

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/DK04/000914

International filing date: 23 December 2004 (23.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: DK  
Number: PA 2004 01843  
Filing date: 26 November 2004 (26.11.2004)

Date of receipt at the International Bureau: 11 February 2005 (11.02.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



# Kongeriget Danmark

Patent application No.: PA 2004 01843

Date of filing: 26 November 2004

Applicant:  
(Name and address) Aros Applied Biotechnology ApS  
c/o Østjysk Innovation  
Gustav Wieds Vej 10  
DK-8000 Århus C  
Denmark

Title: Classification of cancer

IPC: -

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.



**Patent- og Varemærkestyrelsen**  
Økonomi- og Erhvervsministeriet

03 February 2005

*Susanne Morsing*  
Susanne Morsing



26 NOV. 2004

PVS

**Classification of cancer****Field of invention**

The present invention relates to a method for classification of cancer in an individual, wherein the microsatellite status and a prognostic marker are determined by examining gene expression patterns. The invention also relates to various methods of treatment of cancer. Additionally, the present invention concerns a pharmaceutical composition for treatment of cancer and uses of the present invention. The invention also relates to an assay for classification of cancer.

**Background of invention**

Studies of differential gene expression in diseased and normal tissues have been greatly facilitated by the building of large databases of the human genome sequences. Gene expression alterations are important factors in the progression from normal tissue to diseased tissue. In order to obtain a profile of transcriptional status in a certain cell type or tissue, array-based screening of thousands of genes simultaneously is an invaluable tool. Array-based screening even allows for the identification of key genes that alone, or in combination with other genes, regulate the behaviour of a cell or tissue. Candidate genes for future therapeutic intervention may thus also be identified.

Colorectal cancer generally occurs in 1 out of every 20 individuals at some point during their lifetime. In the United States alone about 150,000 new cases are diagnosed each year which amount to 15% of the total number of new cancer diagnoses. Unfortunately, colorectal cancer causes about 56,000 deaths a year in the United States.

The malignant transformation from normal tissue to cancer is believed to be a multistep process. Two molecular pathways are known to be involved in the development of colorectal cancer (Lengauer C, Kinzler KW, Vogelstein B., 1998) namely the microsatellite stable (MSS) pathway and the microsatellite instable (MSI) pathway. MSS is associated with high frequency of allelic losses, abnormalities of cytogenetic nature and abnormal tumor content of DNA. MSI however is associated with defects in the DNA mismatch repair system which leads to increased rate of point mutations and minor chromosomal insertions or deletions.

## **Classification of cancer**

### **Field of invention**

The present invention relates to a method for classification of cancer in an individual, wherein the microsatellite status and a prognostic marker are determined by examining gene expression patterns. The invention also relates to various methods of treatment of cancer. Additionally, the present invention concerns a pharmaceutical composition for treatment of cancer and uses of the present invention. The invention also relates to an assay for classification of cancer.

### **Background of invention**

Studies of differential gene expression in diseased and normal tissues have been greatly facilitated by the building of large databases of the human genome sequences. Gene expression alterations are important factors in the progression from normal tissue to diseased tissue. In order to obtain a profile of transcriptional status in a certain cell type or tissue, array-based screening of thousands of genes simultaneously is an invaluable tool. Array-based screening even allows for the identification of key genes that alone, or in combination with other genes, regulate the behaviour of a cell or tissue. Candidate genes for future therapeutic intervention may thus also be identified.

Colorectal cancer generally occurs in 1 out of every 20 individuals at some point during their lifetime. In the United States alone about 150,000 new cases are diagnosed each year which amount to 15% of the total number of new cancer diagnoses. Unfortunately, colorectal cancer causes about 56,000 deaths a year in the United States.

The malignant transformation from normal tissue to cancer is believed to be a multistep process. Two molecular pathways are known to be involved in the development of colorectal cancer (Lengauer C, Kinzler KW, Vogelstein B., 1998) namely the microsatellite stable (MSS) pathway and the microsatellite instable (MSI) pathway. MSS is associated with high frequency of allelic losses, abnormalities of cytogenetic nature and abnormal tumor content of DNA. MSI however is associated with defects in the DNA mismatch repair system which leads to increased rate of point mutations and minor chromosomal insertions or deletions.

MSI tumors can be of hereditary or sporadic nature. Ninety percent of MSI tumours are of sporadic origin. Sporadic tumours are presumably MSI due to epigenetic hypermethylation of the MLH1 gene promoter. The hereditary tumours account for 10 % of the MSI tumors. Mutations of for example the MLH1 or MSH 2 genes are often the cause of hereditary tumor development.

The ability of being able to determine the sporadic or hereditary nature of a MSI tumor is highly valuable. In case a tumor is characterized as being MSI , and certain clinical criteria are fulfilled such as age below 50 or three first degree relatives with colon cancer, a screening programme of family members for early diagnosis and treatment of potential colon or endometrial cancer development is initiated. The human and economic costs in relation to screening programmes are severe. Consequently, a need for identifying colon cancers with a hereditary character exists. Further, these patients have a poor prognosis, as they have an increased risk of metachronous colon tumors and a highly increased risk of getting cancer in the endometrium (females), upper urinary tract and a number of other organs. Thus, one may regard the determination of a colon tumor as being sporadic or hereditary as determination of a prognostic factor.

Tumors appearing to be similar – morphologically, histochemically or microscopically – can be profoundly different. They can have different invasive and metastasizing properties, as well as respond differently to therapy. There is thus a need in the art for methods which distinguish tumors and tissues on different bases than are currently in use in the clinic. Determination of microsatellite status using an array-based methodology is faster than conventional DNA based methods, as it does not require microdissection, and forms a set of genes that can be combined with other sets of genes on a colon cancer array that can be used to determine microsatellite status as well as e.g. predict disease course by identifying hereditary cases or other prognostic important factors, and finally predict therapy response.

### **Summary of invention**

In one aspect the present invention relates to a method of classifying cancer in an individual having contracted cancer comprising

in a sample from the individual having contracted cancer determining the microsatellite status of the tumor and

5 in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount which forms a pattern, determining from said pattern a prognostic marker, wherein the microsatellite status and the prognostic marker is determined simultaneously or sequentially

classifying said cancer from the microsatellite status and the prognostic marker.

10 The cancer may be any cancer known to be microsatellite instable in at least a fraction of the cases, such as colon cancer, uterine cancer, ovary cancer, stomach cancer, cancer in the small intestine, cancer in the biliary system, urinary tract cancer, brain cancer or skin cancer. These cancers are part of the spectrum of cancers that  
15 belong to the hereditary non-polyposis colon cancer syndrome, but the invention is not limited to this syndrome.

Gene expression patterns may be formed by only a few genes, but it is also a preferred embodiment that a multiplicity of genes form the expression pattern whereby  
20 information for classification of cancer can be obtained.

Furthermore, the invention relates to a method for classification of cancer in an individual having contracted cancer, wherein the microsatellite status is determined by a method comprising the steps of

25 in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the microsatellite status of said cancer,

30 determining the presence and/or amount of said gene expression products forming said pattern,

obtaining an indication of the microsatellite status of said cancer in the individual based on the step above.

35

Yet another aspect of the invention relates to a method for classification cancer in an individual having contracted cancer, wherein the hereditary or sporadic nature is determined by a method comprising the steps of

5 in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the hereditary or sporadic nature of said cancer,

10 determining the presence and/or amount of said gene expression products forming said pattern,

obtaining an indication of the hereditary or sporadic nature of said cancer in the individual based on the step above.

15 The present invention further concerns a method for treatment of an individual comprising the steps of

20 selecting an individual having contracted a colon cancer, wherein the microsatellite status is stable, determined according to any of the methods as defined herein

treating the individual with anti cancer drugs .

25 Another aspect of the present invention relates to a method for treatment of an individual comprising the steps of

selecting an individual having contracted a colon cancer, wherein the microsatellite status is instable, determined according to any of the methods as defined herein

30 treating the individual with anti cancer drugs.

Yet another aspect of the present invention relates to a method for reducing malignancy of a cell, said method comprising

contacting a tumor cell in question with at least one peptide expressed by at least one gene selected from genes being expressed at least two-fold higher in tumor cells than the amount expressed in said tumor cell in question.

5 Additionally, the present invention concerns a method for reducing malignancy of a tumor cell in question comprising,

obtaining at least one gene selected from genes being expressed at least two fold lower in tumor cells than the amount expressed in normal cells

10 introducing said at least one gene into the tumor cell in question in a manner allowing expression of said gene(s).

The invention also relates to a method for reducing malignancy of a cell in question, said method comprising

15 obtaining at least one nucleotide probe capable of hybridising with at least one gene of a tumor cell in question, said at least one gene being selected from genes being expressed in an amount at least two-fold higher in tumor cells than the amount expressed in normal cells, and

20 introducing said at least one nucleotide probe into the tumor cell in question in a manner allowing the probe to hybridise to the at least one gene, thereby inhibiting expression of said at least one gene.

25 In a further aspect the invention relates to a method for producing antibodies against an expression product of a cell from a biological tissue, said method comprising the steps of

30 obtaining expression product(s) from at least one gene said gene being expressed as defined herein

immunising a mammal with said expression product(s) obtaining antibodies against the expression product.

35

The present invention also concerns a method for treatment of an individual comprising the steps of

5 selecting an individual having contracted a colon cancer, wherein the microsatellite status is stable, determined according to any of the methods as defined herein

introducing at least one gene into the tumor cell in a manner allowing expression of said gene(s).

10 The present invention further relates to a pharmaceutical composition for the treatment of a classified cancer comprising at least one antibody as defined herein.

15 In yet another aspect the invention concerns a pharmaceutical composition for the treatment of a classified cancer comprising at least one polypeptide as defined herein.

Further, the invention relates to a pharmaceutical composition for the treatment of a classified cancer comprising at least one nucleic acid and/or probe as defined herein.

20 In an additional aspect the present invention relates to an assay for classification of cancer in an individual having contracted cancer, comprising

25 at least one marker capable of determining the microsatellite status in a sample and

at least one marker in a sample determining the prognostic marker, wherein the microsatellite status and the prognostic marker is determined simultaneously or sequentially.

### 30 Detailed description of the drawings

#### Figure 1

**Unsupervised hierarchical clustering of colorectal tumors based on the 1239 genes with the highest variation across all tumors.**

35 The phylogenetic tree shows the spontaneous clustering of tumor samples and normal biopsies. Germline mutation indicates samples with hereditary mutations in

either *MLH1* or *MSH2* genes. In columns referring to results of immunohistochemistry a plus indicates a positive antibody staining. Tumor location indicates right-sided or left-sided location in the colon of the tumor.

5 **Figure 2**

**Summary of the performance of the microsatellite instability classifier based on microarray data.**

Panel A shows the number of classification errors as a function of the number of genes used. Panel B shows  $\log_2$  of the ratio of the distance between a tumor to the centers of the microsatellite instable group and the microsatellite stable tumors. A value of +2 indicates that the distance of a tumor to the microsatellite instable group is 4 times the distance to the microsatellite stable group. Open bars are MSI tumors and solid bars are MSS tumors. Panel C shows the result of the permutation analysis for estimation of the stability of the classifier. This was estimated by generating one hundred new classifiers based on randomly chosen datasets from the 101 tumors each consisting of 30 microsatellite stable and 25 microsatellite instable samples. In each case the classifier was tested with the remaining 46 samples. The performance for each set was evaluated and averaged over all 100 training and test sets.

20

**Figure 3**

**Classification of MSI tumors as hereditary or sporadic cases based on two genes.**

Panel A shows the number of classification errors as a function of the number of genes used. In crossvalidation we found a minimum number of one error using two genes and adding more genes increased the number of errors to a maximum number of twelve. Both genes were used in at least 36 of the 37 crossvalidation loops. Panel B shows  $\log_2$  of the ratio of the distance between a tumor to the centers of the sporadic microsatellite instable group and the hereditary microsatellite instable group. Panel C shows microarray signal values for *MLH1* and *PIWIL1* genes for all tumors. Asterisk indicates the misclassified tumor

30

35



**Figure 4****Classification of microsatellite-instability status based on real-time PCR.**

Panel A shows a cluster analysis of 18 of the 101 tumors samples and 9 genes based on the microarray data and compared to real-time PCR data from same samples and genes. Dark colors indicate relative low expression and light/light grey color palette high expression. Panel B shows the result of 47 new independent samples based on PCR data from 7 of the 9 genes. Relative distances are explained in the legend to figure 2. The two misclassified tumors are indicated with an asterisk. For PCR primers and hybridization probes see supplement to methods.

**Figure 5**

Kaplan-Meier estimates of crude survival among patient with Stage II and Stage III colorectal cancer according to microsatellite status of the tumor, determined by gene expression. Open triangles indicate censored samples. The patients left at risk are denoted in brackets. The P values were calculated with use of the log-rank test.

**Figure 6**

Phylogenetic tree resulting from unsupervised hierarchical clustering. Clusteranalysis of colon specimens with associated clinicopathological features.

**Figure 7**

Multidimensional scaling plot showing distances between groups of tumors.

**Figure 8**

Performance of prediction of survival before and after separation in MSI-H and MSS

**Figure 9**

Performance of the classifier for identification of hereditary disease.

**Figure 10**

Kaplan Meier estimates of overall survival among patients with Dukes' B and Dukes' C colon cancer according to microsatellite-instability status of the tumor, determined by gene expression.

**Detailed description of the invention****Classification of cancer**

The present inventors have, using large-scale array-based screenings, found a pool of genes, the expression products of which may be used to classify cancer in an individual. The presence of expression products and level of expression products provides an expression pattern which is correlated to a specific status and/or prognostic marker of the cancer. Characterization of the genes or functional analysis of the gene expression products as such is not required to classify the cancer based on the present method. Thus, the expression products of the plurality of genes can be used as markers for the classification of disease.

One aspect of the present invention concerns a method for classifying cancer in an individual having contracted cancer by determining the microsatellite status and a prognostic marker in a sample. Determination of the microsatellite status and the prognostic marker may be performed simultaneously or sequentially. In one embodiment of the present invention the microsatellite status is determined. The prognostic marker is determined in a sample, wherein the presence and/or the amount of a number of gene expression products form a pattern wherefrom the prognostic marker is determined. Based on the information gathered from the microsatellite status and the prognostic marker the cancer can be classified. In a preferred embodiment the prognostic marker is the hereditary or sporadic nature of the cancer. The hereditary or sporadic nature of the cancer can be determined through a number of steps comprising determining the presence and/or amount of gene expression products forming a pattern in a sample. The sample comprises a number of gene expression products the presence and/or amount of which forms a pattern that is indicative of the hereditary or sporadic nature of the cancer. Hereby, an indication of the hereditary or sporadic nature of the cancer is obtained.

In one embodiment of the invention the microsatellite status is determined using conventional analysis of microsatellite status as described elsewhere herein.

In another embodiment of the present invention the microsatellite status is determined by gene expression patterns wherein the presence and/or the amount of the gene expression products form a pattern that is indicative of the microsatellite status.

Classification of cancer provides knowledge of the survival chances of an individual having contracted cancer. In case of cancer which according to the present invention has been classified as a hereditary cancer, screening programmes of family members to the individual having the classified cancer can be initiated. Such  
5 screening programmes can comprise conventional screening programmes employing sequencing and other methods as described elsewhere. Thus, individuals at risk of developing cancer may be identified and action taken accordingly to detect developing cancer at an early stage of the disease greatly improving the chances of successful intervention and thus survival rates.

10 Classification of cancer also provides insights on which sort of treatment should be offered to the individual having contracted cancer, thus providing an improved treatment response of the individual. Likewise, the individual may be spared treatment that is inefficient in treating the particular class of cancer and thus spare the  
15 individual severe side effects associated with treatment that may even not be suitable for the class of cancer.

#### **Microsatellite status**

The use of highly variable repetitive sequences found in microsatellite regions adjacent to genes or other areas of interest may be used as markers for linkage analysis, DNA fingerprinting, or other diagnostic application.  
20

Microsatellites are defined as loci (or regions within DNA sequences) where short sequences of DNA are repeated in tandem repeats. This means that the sequences  
25 are repeated one right after the other. The lengths of sequences used most often are di-, tri-, or tetra-nucleotides. At the same location within the genomic DNA the number of times the sequence (ex. AC) is repeated often varies between individuals, within populations, and/or between species. Due to the many repeats the microsatellites are prone to alter if there is a reduced repair of mismatches in the genome. In  
30 the present invention the traditional method of determining microsatellite status by employing microsatellite markers is replaced by determination of gene expression patterns.

An important factor in multi-step carcinogenesis is genomic instability. The development of some cancer forms is known to follow two distinct molecular routes. One  
35

route is the microsatellite stable, MSS, (and chromosomal instable pathway) which is often associated with a high frequency of allelic losses, cytogenetic abnormalities and abnormal DNA tumor contents. The second route is the microsatellite instable pathway MSI that is characterized by defects in the DNA mismatch repair system which leads to a high rate of point mutations and small chromosomal insertions and deletions. The small chromosomal insertions and deletions can be detected as mono and dinucleotide repeats (Boland CR, Thibodeau SN, Hamilton SR, et al., Cancer Res 1998;58(22):5248-57).

One aspect of the present invention relates to the classification of cancer in an individual having contracted cancer by determining the microsatellite status and a prognostic marker. One embodiment of the invention relates to microsatellite status determined by conventional methods employing microsatellite analysis as described above. Another embodiment of the invention relates to establishing the microsatellite status by determining the presence and/or amount of gene expression products of a sample which comprises a plurality of gene expression products forming a pattern which is indicative of the microsatellite status.

The expression products of genes according to the present invention are not necessarily identical to the genes that are analysed by microsatellite markers in conventional methods of determining microsatellite status. The pattern of the gene expression products according to the present invention however correlates with information on microsatellite status that can be obtained using traditional methods.

The determination of the microsatellite status and the prognostic marker of the cancer may be performed sequentially. However, the determinations may also be performed simultaneously.

#### **Prognostic marker**

Together with knowledge of the microsatellite status in a sample of an individual having contracted cancer a prognostic marker is employed for classifying the cancer. The prognostic marker may be any marker that provides knowledge of the cancer type when combined with knowledge of microsatellite status. Consequently the prognostic marker may provide additional information on the cancer type when the microsatellite status is stable and similarly when the microsatellite status is

instable. In a preferred embodiment of the present invention the prognostic marker is the hereditary or sporadic nature of a cancer given that the microsatellite status is instable. The prognostic marker may in another embodiment be a prognostic marker for any feature or trait that provides further possibilities of classifying cancer.

- 5 The prognostic marker is determined in a sample comprising a number of gene expression products wherein the presence and/or amounts of gene expression products form a pattern that is indicative of the prognostic marker.

#### **Hereditary and sporadic nature of cancer**

- 10 Hereditary nonpolyposis colon cancer (HNPCC) is a hereditary cancer syndrome which carries a very high risk of colon cancer and an above-normal risk of other cancers (uterus, ovary, stomach, small intestine, biliary system, urinary tract, brain, and skin). The HNPCC syndrome is due to mutation in a gene in the DNA mismatch repair system, usually the MLH1 or MSH2 gene or less often the MSH6 or PMS2
- 15 genes. Families with HNPCC account for about 5% of all cases of colon cancer and typically have the following features (called the Amsterdam clinical criteria):

- Three or more first relative family members with colorectal cancer; affected family members in two or more generations; and at least one person with colon cancer
- 20 diagnosed before the age of 50.

- The highest risk with HNPCC is for colon cancer. A person with HNPCC has about an 80% lifetime risk of colon cancer. Two-thirds of these tumors occur in the proximal colon. Women with HNPCC have a 20-60% lifetime risk of endometrial cancer.
- 25 In HNPCC, the gastric cancer is usually intestinal-type adenocarcinoma. The ovarian cancer in HNPCC may be diagnosed before age 40. Other HNPCC-related cancers have characteristic features: the urinary tract cancers are transitional carcinoma of the ureter and renal pelvis; the small bowel cancer is most common in the duodenum and jejunum; and the most common type of brain tumor is glioblastoma.
- 30 The diagnosis of HNPCC may be made on the basis of the Amsterdam clinical criteria (listed above) or on the basis of molecular genetic testing for mutations in a mismatch repair gene (MLH1, MSH2, MSH6 or PMS2). Mutations in MLH1 and MSH2 account for 90% of HNPCC. Mutations in MSH6 and PMS2 account for the rest.

HNPCC is inherited in an autosomal dominant manner. Each child of an individual with HNPCC has a 50% chance of inheriting the mutation. Most people diagnosed with HNPCC have inherited the condition from a parent. However, not all individuals with an HNPCC gene mutation have a parent who had cancer. Prenatal diagnosis for pregnancies at increased risk for HNPCC is possible.

In tumors that are microsatellite instable it is often found that the DNA mismatch repair proteins that are encoded by the *MLH1* or *MSH2* genes are inactivated. In case of microsatellite instable hereditary non-polyposis colorectal cancers germline mutation in *MLH1* and *MSH2* and somatic loss of function of the normal allele has been found to be associated with the disease.

For most sporadic MSI tumors epigenetic hypermethylation of the *MLH1* promoter can be found to be associated with the cancer (Cunningham JM, Christensen ER, Tester DJ, et al., *Cancer Res* 1998;58(15):3455-60., Kane MF, Loda M, Gaida GM, et al., *Cancer Res* 1997;57(5):808-11., Herman JG, Umar A, Polyak K, et al., *Proc Natl Acad Sci U S A* 1998;95(12):6870-5., Kuismanen SA, Holmberg MT, Salovaara R, de la Chapelle A, Peltomaki P., *Am J Pathol* 2000;156(5):1773-9).

## 20 **Forms of cancer**

Cancer leads to a change in the expression of one or more genes. The methods according to the invention may be used for classifying cancer according to the microsatellite status and/or the hereditary or sporadic nature of the cancer. Thus, the cancer may be any malignant condition in which genomic instability is involved in the development of cancer, such as cancers related to hereditary non-polyposis colorectal cancer, such as endometrial cancer, gastric cancer, small bowel cancer, ovarian cancer, kidney cancer, pelvic renal cancer or tumors of the nervous system, such as glioblastoma.

30 One particular form of cancer according to the present invention is that of the colon/rectum.

The cancer may be of any tumor type, such as an adenocarcinoma, a carcinoma, a teratoma, a sarcoma, and/or a lymphoma.

35

In relation to the gastro-intestinal tract, the biological condition may also be colitis ulcerosa, Mb. Crohn, diverticulitis, adenomas.

### **Colorectal tumors**

5 The data presented herein relates to colorectal tumors and therefore the description has focused on the gene expression level as one manner of identifying genes involved in the prediction of survival in cancer tissue. The malignant progression of cancer of colon or rectum may be described using Dukes stages where normal mu-  
cosa may progress to Dukes A superficial tumors to Dukes B, slightly invasive tu-  
10 mors, to Dukes C that have spread to lymphnodes and finally to Dukes D that have metastasized to other organs.

The grade of a tumor can also be expressed on a scale of I-IV. The grade reflects the cytological appearance of the cells. Grade I cells are almost normal, whereas  
15 grade II cells deviate slightly from normal. Grade III appear clearly abnormal, whereas grade IV cells are highly abnormal.

The phrase colon cancer is in this application meant to be equivalent to the phrase colorectal cancer. Colon cancers may be located in the right side of the colon, the  
20 left side of the colon, the transverse part of the colon and/or in the rectum.

### **Samples**

The samples according to the present invention may be any cancer tissue.  
The sample may be in a form suitable to allow analysis by the skilled artisan, such  
25 as a biopsy of the tissue, or a superficial sample scraped from the tissue. In one embodiment of the invention it is preferred that the sample is from a resected colon cancer tumor. In another embodiment the sample may be prepared by forming a suspension of cells made from the tissue. The sample may, however, also be an extract obtained from the tissue or obtained from a cell suspension made from the  
30 tissue. The sample may be fresh or frozen, or treated with chemicals.

### **Expression pattern**

Expression of one gene or more genes in a sample forms a pattern that is character-  
istic of the state of the cell. In a sample from an individual having contracted cancer  
35 a plurality of gene expression products are present. By expression pattern is meant

the presence of a combination of a number of expression products and/or the amount of expression products specific for a given biological condition, such as cancer. The pattern is produced by determining the expression products of selected genes that together reveals a pattern that is indicative of the biological condition.

5 Thus, a selection of the genes that carry information about a specific condition is developed. Selection of the genes is achieved by analyzing large numbers of genes and their expression products to find the genes that will enable the desired differentiation between various conditions, such as microsatellite status (MSS or MSI) and/or prognostic marker, such as for example the sporadic or hereditary nature of a  
10 given cancer sample. The criteria for selection of the best genes for the pattern to be indicative of given biological conditions include confidence levels i.e. how accurate are the selected genes forming an expression pattern in giving correct information of the biological condition. Thus, in one aspect of the present invention a specific pattern of gene expression profiles can be used to determine the microsatellite status in  
15 the sample. In a second aspect of the present invention the microsatellite status is determined and a specific pattern of the presence of a plurality of gene expression products and/or amount wherefrom a prognostic marker is determined.

#### **Determination of the microsatellite status employing gene expression patterns**

20 One aspect of the invention specifically relates to a method for determining the microsatellite status in a sample of an individual having contracted cancer based on determination of the expression pattern of at least two genes, such as at least three genes, such as at least four genes, such as at least 5 genes, such as at least 6 genes, such as at least 7 genes, such as at least 8 genes, such as at least 9 genes,  
25 such as at least 10 genes, such as at least 15 genes, such as at least 20 genes, such as at least 30 genes, such as at least 40 genes, such as at least 50 genes, such as at least 60 genes, such as at least 70 genes, such as at least 80 genes, such as at least 90 genes, such as at least 126 genes selected from the group of genes listed in Table 1 below

30 Table 1

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
chemokine (C-C motif) ligand 5	<u>NM_002985</u>	CCL5	1	
tryptophanyl-tRNA synthetase	<u>NM_004184</u>	WARS	2	
proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)	NM_006263	PSME1	3	
bone marrow stromal cell antigen 2	<u>NM_004335</u>	BST2	4	
ubiquitin-conjugating enzyme E2L 6	NM_004223	UBE2L6	5	



A kinase (PRKA) anchor protein 1	NM_003488	AKAP1	6
proteasome (prosome, macropain) activator subunit 2 (PA28 beta)	NM_002818	PSME2	7
carcinoembryonic antigen-related cell adhesion molecule 5	NM_004363	CEACAM5	8
FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)	NM_005766	FARP1	9
myosin X	NM_012334	MYO10	10
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11
autocrine motility factor receptor	NM_001144	AMFR	12
dimethylarginine dimethylaminohydrolase 2	NM_013974	DDAH2	13
tumor necrosis factor, alpha-induced protein 2	NM_006291	TNFAIP2	14
mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	NM_000249	MLH1	15
thymidylate synthetase	NM_001071	TYMS	16
intercellular adhesion molecule 1 (CD54), human	NM_000201	ICAM1	17
rhinovirus receptor			
general transcription factor IIA, 2, 12kDa	NM_004492	GTF2A2	18
Rho-associated, coiled-coil containing protein kinase 2	NM_004850	ROCK2	19
ATP binding protein associated with cell differentiation	NM_005783	TXNDC9	20
NCK adaptor protein 2	NM_003581	NCK2	21
phytanoyl-CoA hydroxylase (Refsum disease)	NM_006214	PHYH	22
metastasis-associated gene family, member 2	NM_004739	MTA2	23
amiloride binding protein 1 (amine oxidase (copper-containing))	NM_001091	ABP1	24
biliverdin reductase A	NM_000712	BLVRA	25
phospholipase C, beta 4	NM_000933	PLCB4	26
chemokine (C-X-C motif) ligand 9	NM_002416	CXCL9	27
purine-rich element binding protein A	NM_005859	PURA	28
quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))	NM_014298	QPRT	29
retinoic acid receptor responder (tazarotene induced) 3	NM_004585	RARRES3	30
chemokine (C-C motif) ligand 4	NM_002984	CCL4	31
forkhead box O3A	NM_001455	FOXO3A	32
interferon, alpha-inducible protein (clone IFI-6-16)	NM_002038	G1P3	34
	NM_022873		123
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35
	NM_005950	MT1G	36
metallothionein 1G	NM_005950		
	NM_000043	TNFRSF6	37
tumor necrosis factor receptor superfamily, member 6	NM_152877		133
	NM_152876		132
	NM_152875		134
	NM_152872		130
	NM_152873		33
	NM_152871		129
	NM_152874		131
endothelial cell growth factor 1 (platelet-derived)	NM_001953	ECGF1	38
SCO cytochrome oxidase deficient homolog 2 (yeast)	NM_005138	SCO2	39
chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant)	NM_006419	CXCL13	40

Granulysin	NM_006433	GNLY	41
CD2 antigen (p50), sheep red blood cell receptor	<u>NM_001767</u>	CD2	42
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43
teratocarcinoma-derived growth factor 1	<u>NM_003212</u>	TDGF1	44
metallothionein 1H	<u>NM_005951</u>	MT1H	45
cytochrome P450, family 2, subfamily B, polypeptide 6	<u>NM_000767</u>	CYP2B6	46
tumor necrosis factor (ligand) superfamily, member 9	<u>NM_003811</u>	TNFSF9	47
	NM_006047	RBM12	48
RNA binding motif protein 12	NM_006047		
heat shock 105kDa/110kDa protein 1	<u>NM_006644</u>	HSPH1	49
staufen, RNA binding protein (Drosophila)	NM_004602	STAU	50
	NM_017452		125
	NM_017453		126
lymphocyte antigen 6 complex, locus G6D	<u>NM_021246</u>	LY6G6D	51
calcium binding protein P22	<u>NM_007236</u>	CHP	52
CDC14 cell division cycle 14 homolog B (S. cerevisiae)	<u>NM_003671</u>	CDC14B	53
	<u>NM_033331</u>		115
epiplakin 1	XM_372063	EPPK1	54
metallothionein 1X	<u>NM_005952</u>	MT1X	55
transforming growth factor, beta receptor II (70/80kDa)	<u>NM_003242</u>	TGFBR2	56
protein kinase C binding protein 1	NM_012408	PRKCBP1	57
	NM_183047		124
transmembrane 4 superfamily member 6	<u>NM_003270</u>	TM4SF6	58
pleckstrin homology domain containing, family B (evectins) member 1	<u>NM_021200</u>	PLEKHB1	59
apolipoprotein L, 1	NM_003661	APOL1	60
	NM_145343		120
indoleamine-pyrrole 2,3 dioxygenase	<u>NM_002164</u>	INDO	61
forkhead box A2	<u>NM_021784</u>	FOXA2	62
granzyme H (cathepsin G-like 2, protein h-CCPX)	<u>NM_033423</u>	GZMH	63
baculoviral IAP repeat-containing 3	NM_001165	BIRC3	64
Homo sapiens metallothionein 1H-like protein		AF333388 (Hs 382039)	135
KIAA0182 protein	<u>NM_014615</u>	KIAA0182	117
G protein-coupled receptor 56	<u>NM_005682</u>	GPR56	65

	<u>NM_201524</u>		116
metallothionein 2A	<u>NM_005953</u>	MT2A	66
F-box only protein 21	NM_015002	FBXO21	67
	NM_012156,	EPB41L1	68
erythrocyte membrane protein band 4.1-like 1	<u>NM_012156</u>		
hypothetical protein MGC21416	<u>NM_173834</u>	MGC21416	69
protein O-fucosyltransferase 1	<u>NM_015352</u> ,	POFUT1	70
	<u>NM_015352</u>		
metallothionein 1E (functional)	<u>NM_175617</u>	MT1E	71
troponin T1, skeletal, slow	<u>NM_003283</u>	TNNT1	72
chimerin (chimaerin) 2	<u>NM_004067</u>	CHN2	73
heterogeneous nuclear ribonucleoprotein H1 (H)	<u>NM_005520</u>	HNRPH1	74
ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle	<u>NM_004046</u>	ATP5A1	75
eukaryotic translation initiation factor 5A	<u>NM_001970</u>	EIF5A	76
perforin 1 (pore forming protein)	<u>NM_005041</u>	PRF1	77
OGT(O-GlcNAc transferase)-interacting protein 106 KDa	<u>NM_014965</u>	OIP106	78
DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	<u>NM_017895</u>	DDX27	79
vacuolar protein sorting 35 (yeast)	<u>NM_018206</u>	VPS35	80
tripartite motif-containing 44	<u>NM_017583</u>	TRIM44	81
transmembrane, prostate androgen induced RNA	NM_020182	TMEPAI	82
	NM_199169		127
	NM_199170		128
dynein, cytoplasmic, light polypeptide 2A	NM_014183	DNCL2A	83
	NM_177953		122
leucine aminopeptidase 3	<u>NM_015907</u>	LAP3	84
chromosome 20 open reading frame 35	<u>NM_018478</u>	C20orf35	85
	NM_033542		118
solute carrier family 38, member 1	<u>NM_030674</u>	SLC38A1	86
CGI-85 protein	NM_016028	CGI-85	87
death associated transcription factor 1	NM_022105,	DATF1	88
	NM_080796		121
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
sestrin 1	<u>NM_014454</u>	SESN1	90
hypothetical protein FLJ20315	<u>NM_017763</u>	FLJ20315	91
hypothetical protein FLJ20647	<u>NM_017918</u>	FLJ20647	92
membrane protein expressed in epithelial-like lung adenocarcinoma	<u>NM_024792</u>	CT120	93
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide	<u>NM_014314</u>	RIG-I	94
keratin 23 (histone deacetylase inducible)	<u>NM_015515</u> ,	KRT23	95
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-	<u>NM_007210</u>	GALNT6	96

acetylgalactosaminyltransferase 6 (GalNAc-T6)	<u>NM_020183</u>	ARNTL2	97
aryl hydrocarbon receptor nuclear translocator-like 2			
apobec-1 complementation factor	NM_014576,	ACF	98
	NM_138932		119
hypothetical protein FLJ20232	<u>NM_019008</u>	FLJ20232	99
apolipoprotein L, 2	NM_030882,	APOL2	100
	NM_145343		120
mitochondrial solute carrier protein	NM_016612	MSCP	101
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
	<u>NM_003011</u> ,		103
SET translocation (myeloid leukaemia-associated)	<u>1</u>	SET	
	<u>Xm_030577</u> ,		104
ATPase, class II, type 9a	<u>9</u>	ATP9a	

One embodiment of the invention concerning the determination of microsatellite status is based on the expression pattern of at least 2 genes, such as at least 3 genes, such as at least 4 genes, such as at least 5 genes, such as at least 6 genes, such as at least 7 genes, such as at least 8 genes, such as at least 9 genes, such as at least 10 genes, such as at least 15 genes, such as at least 20 genes, such as at least 25 genes selected from the group of genes listed in Table 2.

Table 2

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
chemokine (C-C motif) ligand 5	<u>NM_002985</u>	CCL5	1	
tryptophanyl-tRNA synthetase	<u>NM_004184</u>	WARS	2	
proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)	NM_006263	PSME1	3	
bone marrow stromal cell antigen 2	<u>NM_004335</u>	BST2	4	
ubiquitin-conjugating enzyme E2L 6	NM_004223	UBE2L6	5	
A kinase (PRKA) anchor protein 1	NM_003488	AKAP1	6	
proteasome (prosome, macropain) activator subunit 2 (PA28 beta)	<u>NM_002818</u>	PSME2	7	
carcinoembryonic antigen-related cell adhesion molecule 5	<u>NM_004363</u>	CEACAM5	8	
FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)	<u>NM_005766</u>	FARP1	9	
myosin X	<u>NM_012334</u>	MYO10	10	
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
autocrine motility factor receptor	NM_001144	AMFR	12	
dimethylarginine dimethylaminohydrolase 2	<u>NM_013974</u>	DDAH2	13	
tumor necrosis factor, alpha-induced protein 2	<u>NM_006291</u>	TNFAIP2	14	
mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	<u>NM_000249</u>	MLH1	15	
thymidylate synthetase	<u>NM_001071</u>	TYMS	16	

intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	<u>NM_000201</u>	ICAM1	17
general transcription factor IIA, 2, 12kDa	<u>NM_004492</u>	GTF2A2	18
Rho-associated, coiled-coil containing protein kinase 2	<u>NM_004850</u>	ROCK2	19
ATP binding protein associated with cell differentiation	<u>NM_005783</u>	APACD	20
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43
protein kinase C binding protein 1	<u>NM_012408</u>	PRKCBP1	57
	<u>NM_183047</u>		124
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104

or from

## 5 Table 3

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
NCK adaptor protein 2	<u>NM_003581</u>	NCK2	21	
phytanoyl-CoA hydroxylase (Refsum disease)	<u>NM_006214</u>	PHYH	22	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
amiloride binding protein 1 (amine oxidase (copper-containing))	<u>NM_001091</u>	ABP1	24	
biliverdin reductase A	<u>NM_000712</u>	BLVRA	25	
phospholipase C, beta 4	<u>NM_000933</u>	PLCB4	26	
chemokine (C-X-C motif) ligand 9	<u>NM_002416</u>	CXCL9	27	
purine-rich element binding protein A	<u>NM_005859</u>	PURA	28	
quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))	<u>NM_014298</u>	QPRT	29	
retinoic acid receptor responder (tazarotene induced) 3	<u>NM_004585</u>	RARRES3	30	
chemokine (C-C motif) ligand 4	<u>NM_002984</u>	CCL4	31	
forkhead box O3A	<u>NM_001455</u>	FOXO3A	32	
metallothionein 1X	<u>NM_005952</u>	MT1X	55	
interferon, alpha-inducible protein (clone IFI-6-16)	<u>NM_002038</u>	G1P3	34	
	<u>NM_022873</u>		123	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
	<u>NM_005950</u>	MT1G	36	
metallothionein 1G	<u>NM_005950</u>			

tumor necrosis factor receptor superfamily, member 6	NM_000043	TNFRSF6	37
	NM_152877		133
	NM_152876		132
	NM_152875		134
	NM_152872		130
	NM_152873		33
	NM_152871		129
endothelial cell growth factor 1 (platelet- derived)	NM_152874		131
	NM_001953	ECGF1	38
SCO cytochrome oxidase deficient homolog 2 (yeast)	NM_005138	SCO2	39
chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant)	NM_006419	CXCL13	40
	NM_006433	GNLY	41
Granulysin splicing factor, arginine/serine-rich 6	NM_006275	SFRS6	43
	NM_012408	PRKCBP1	57
	NM_183047		124
protein kinase C binding protein 1			
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89
hypothetical protein FLJ20618	NM_017903	FLJ20618	102
SET translocation (myeloid leukaemia- associated)	NM_003011.1	SET	103
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104

or from

## 5 Table 4

Gene name	Ref seq	Gene sym- bol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11	
metastasis-associated gene family, member 2	NM_004739	MTA2	23	
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35	
CD2 antigen (p50), sheep red blood cell recep- tor	NM_001767	CD2	42	
splicing factor, arginine/serine-rich 6	NM_006275	SFRS6	43	
teratocarcinoma-derived growth factor 1	NM_003212	TDGF1	44	
metallothionein 1H	NM_005951	MT1H	45	
cytochrome P450, family 2, subfamily B, poly- peptide 6	NM_000767	CYP2B6	46	
tumor necrosis factor (ligand) superfamily, member 9	NM_003811	TNFSF9	47	
	NM_006047,	RBM12	48	
RNA binding motif protein 12	NM_006047			
heat shock 105kDa/110kDa protein 1	NM_006644	HSPH1	49	
staufen, RNA binding protein (Drosophila)	NM_004602	STAU	50	
	NM_017452		125	
	NM_017453		126	
lymphocyte antigen 6 complex, locus G6D	NM_021246	LY6G6D	51	
calcium binding protein P22	NM_007236	CHP	52	

CDC14 cell division cycle 14 homolog B (S. cerevisiae)	NM_003671 NM_033331	CDC14B	53 115
epiplakin 1	XM_372063	EPPK1	54
metallothionein 1X	NM_005952	MT1X	55
transforming growth factor, beta receptor II (70/80kDa)	NM_003242	TGFR2	56
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 129
transmembrane 4 superfamily member 6	NM_003270	TM4SF6	58
pleckstrin homology domain containing, family B (evectins) member 1	NM_021200	PLEKHB1	59
apolipoprotein L, 1	NM_003661 NM_145343	APOL1	60 125
indoleamine-pyrrole 2,3 dioxygenase	NM_002164 NM_021784	INDO FOXA2	61 62
forkhead box A2	NM_021784		
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89
mitochondrial solute carrier protein	NM_016612 NM_016612	MSCP	101
hypothetical protein FLJ20618	NM_017903	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	NM_003011.1	SET	103
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104

or from

## 5 Table 5

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11	
metastasis-associated gene family, member 2	NM_004739	MTA2	23	
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35	
splicing factor, arginine/serine-rich 6	NM_006275	SFRS6	43	
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
granzyme H (cathepsin G-like 2, protein h-CCPX)	NM_033423	GZMH	63	
baculoviral IAP repeat-containing 3	NM_001165 NM_001165	BIRC3	64	
Homo sapiens metallothionein 1H-like protein		AF333388 (Hs 382039)	135	
KIAA0182 protein	NM_014615	KIAA0182	117	
G protein-coupled receptor 56	NM_005682 NM_301524	GPR56	65 116	

metallothionein 2A	<u>NM_005953</u>	MT2A	66
F-box only protein 21	<u>NM_015002</u>	FBXO21	67
erythrocyte membrane protein band 4.1-like 1	<u>NM_012156</u>	EPB41L1	68
hypothetical protein MGC21416	<u>NM_173834</u>	MGC21416	69
protein O-fucosyltransferase 1	<u>NM_015352</u>	POFUT1	70
metallothionein 1E (functional)	<u>NM_175617</u>	MT1E	71
	<u>NM_003283</u>	TNNT1	72
troponin T1, skeletal, slow			
chimerin (chimaerin) 2	<u>NM_004067</u>	CHN2	73
heterogeneous nuclear ribonucleoprotein H1 (H)	<u>NM_005520</u>	HNRPH1	74
ATP synthase, H <sup>+</sup> transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle	<u>NM_004046</u>	ATP5A1	75
eukaryotic translation initiation factor 5A	<u>NM_001970</u>	EIF5A	76
perforin 1 (pore forming protein)	<u>NM_005041</u>	PRF1	77
OGT(O-Glc-NAc transferase)-interacting protein 106 KDa	<u>NM_014965</u>	OIP106	78
DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	<u>NM_017895</u>	DDX27	79
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
hypothetical protein FLJ20232	<u>NM_019008</u>	FLJ20232	99
	<u>NM_030882</u> ,	APOL2	100
apolipoprotein L, 2	<u>NM_145343</u>		120
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104

or from

Table 6

5

Gene name	Ref seq	Gene sym- bol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
metallothionein 1G	<u>NM_005950</u>	MT1G	36	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	
protein kinase C binding protein 1	<u>NM_012408</u>	PRKCBP1	57	
	<u>NM_183047</u>		129	
vacuolar protein sorting 35 (yeast)	<u>NM_018206</u>	VPS35	80	
tripartite motif-containing 44	<u>NM_017583</u>	TRIM44	81	



	NM_020182	TMEPAI	82
	NM_199169		127
transmembrane, prostate androgen induced RNA	NM_199170		128
dynein, cytoplasmic, light polypeptide 2A	NM_014183	DNCL2A	83
	NM_177953		122
leucine aminopeptidase 3	NM_015907	LAP3	84
chromosome 20 open reading frame 35	NM_018478	C20orf35	85
	NM_033542		118
solute carrier family 38, member 1	<u>NM_030674</u>	SLC38A1	86
CGI-85 protein	NM_016028	CGI-85	87
death associated transcription factor 1	NM_022105, NM_080796	DATF1	88 121
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
sestrin 1	<u>NM_014454</u>	SESN1	90
hypothetical protein FLJ20315	<u>NM_017763</u>	FLJ20315	91
hypothetical protein FLJ20647	<u>NM_017918</u>	FLJ20647	92
membrane protein expressed in epithelial-like lung adenocarcinoma	<u>NM_024792</u>	CT120	93
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide	<u>NM_014314</u>	RIG-I	94
keratin 23 (histone deacetylase inducible)	NM_015515	KRT23	95
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetyltransferase 6 (GalNAc-T6)	<u>NM_007210</u>	GALNT6	96
aryl hydrocarbon receptor nuclear translocator-like 2	<u>NM_020183</u>	ARNTL2	97
apobec-1 complementation factor	NM_014576 NM_138932	ACF	98 119
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104

Another embodiment of the invention concerning the determination of microsatellite status is based on the expression pattern of at least 2 genes, such as at least 3 genes, such as at least 4 genes, such as at least 5 genes, such as at least 6 genes, such as at least 7 genes, such as at least 8 genes, such as at least 9 genes selected from the group of genes listed in Table 7 below.

Table 7

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	
protein kinase C binding protein 1	NM_012408	PRKCBP1	57	
	NM_183047		124	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103	
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104	

Another embodiment of the invention concerning the determination of microsatellite status is based on the expression pattern of at least 2 genes, such as at least 3 genes, such as at least 4 genes, such as at least 5 genes, such as at least 6 genes, such as at least 7 genes selected from the group of genes listed in Table 8 below.

Table 8

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
Splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	
protein kinase C binding protein 1	NM_012408	PRKCBP1	57	
	NM_183047		124	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	

One embodiment of the invention concerning the determination of microsatellite status is based on the expression pattern of at least one gene, for example two genes, for example 3 genes, such as 4 genes selected from the group of genes listed below in Table 9 and at least one gene, for example two genes, for example 3 genes, such as 4 genes, for example 5 genes selected from the group of genes listed below in Table 10

Table 9

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
Splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

and

Table 10

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103	
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104	

5

Another embodiment relates to the determination of microsatellite status is based on the expression pattern of at least one gene, for example 2 genes, for example 3 genes, such as 4 genes selected from the group of genes that are down regulated in MSS colon cancers compared to MSI colon cancers listed below in Table 11, and at least one gene, for example 2 genes, for example 3 genes, such as 4 genes, for example 5 genes selected from the group of genes that are up-regulated in MSS colon cancers compared to MSI colon cancers listed below in Table 12.

10

Table 11

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
Splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

15

Table 12

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89	
hypothetical protein FLJ20618	NM_017903	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	NM_003011.1	SET	103	
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104	

### Sporadic or hereditary classification using gene expression patterns

- One embodiment of the invention relates to a method of determining the hereditary or sporadic nature of cancer as the prognostic marker in an individual having contracted cancer based on determination of the expression pattern of at least 2 genes, such as at least 3 genes, such as at least 4 genes, such as at least 5 genes, such as at least 6 genes, such as at least 7 genes, such as at least 8 genes, such as at least 9 genes, such as at least 10 genes selected from the group of genes listed in Table 13.

Table 13

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
Homeo box C6	NM_004503	HOXC6	105
Piwi - like 1	NM_004764.2	PIWIL1	106
Mut L homolog 1	NM_00249.2	MLH1	107
Collapsin response mediator protein 1	NM_001313.2	CRMP1	108
Homeo box B2	NM_002145.2	HOXB2	109
Pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyd synthetase)	NM_002860.2	PYCS/ADH18 A1	110
TGFB inducible early growth response	NM_005655.1	TIEG	111
Checkpoint with forkhead and ring finger domains	NM_018223.1	CHFR	112
Hypothetical protein FLJ13842	NM_024645.1	FLJ13842	113
Phosphoprotein regulated by mitogenic pathways	NM_025195.1	C8FW	114

- In a further embodiment of the invention the determination of the hereditary or sporadic nature of cancer is based on the expression pattern of at least 2 genes as listed in Table 14.

Table 14

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
Piwi - like 1	NM_004764.2	PIWIL1	105
Mut L homolog 1	NM_00249.2	MLH1	106

In one specific embodiment the MLH1 gene is down regulated in sporadic disease.  
5 In yet another embodiment the PIWIL1 is expressed in lower amounts in hereditary cancer.

#### Gene names and definition

The genes according to the present invention are identified by their gene name,  
10 gene symbol and a reference sequence number (RefSeq). Furthermore, the genes listed have been assigned a SEQ ID No and the sequence is submitted in the accompanying sequence listing. The reference sequence number refers to the Reference Sequence collection prepared by the the National Center for Biotechnologic Information (NCBI) and where a comprehensive set of sequences for  
15 major research organisms is provided, see <http://www.ncbi.nlm.nih.gov/RefSeq>.

#### Sample preparation

A number of procedures for the isolation of nucleic acids (DNA, RNA, mRNA) from a sample are available and well known to a person skilled in the art. Genomic DNA  
20 may be isolated for detection of mutations of the genome, or for detection of copy number of a gene or a number of other applications which will be appreciated by the skilled artisan. RNA and especially mRNA will be isolated when expression levels of a gene or several genes are to be detected. The sample may be from fresh or frozen tissue as defined elsewhere herein.

25 Before analyzing the sample by for example oligonucleotide arrays or quantitative PCR one or more preparations of the sample may be performed. Often, sample preparations include extraction of intracellular material for example extraction of nucleic acids from whole cell samples, viruses and the like, amplification of nucleic  
30 acids, fragmentation, transcription, labelling and/or extension reactions. One or more of these preparation features may be incorporated readily into the present invention.

### RNA extraction

Methods of isolating total RNA and/or mRNA are well known to those skilled in the art. In one embodiment total RNA is isolated from a given sample by extraction using acidic guanidinium-phenol-chloroform extraction. mRNA may be isolated from RNA or directly for example based on the polyadenylated tail of mRNA using oligo dT column chromatography or by using (dT)<sub>n</sub> magnetic beads (see, e.g. Sambrook et al., *Molecular Cloning: A laboratory Manual* 2<sup>nd</sup> Ed.), Vols 1-3, Cold Spring Harbour Laboratory (1989), or *Current Protocols in Molecular Biology*, F. Ausubel et al., ed. Greene Publishing and Wiley-Interscience, New York (1987).

### PCR

PCR (Polymerase Chain Reaction) is a key technique in molecular genetics that permits the analysis of any short sequence of DNA or RNA without having to clone the short sequence first. PCR is used to reproduce (amplify) selected sections of DNA or RNA. PCR amplification generally involves the use of one strand of the target nucleic acid sequence as a template for producing a large number of complements to that sequence. Generally, two primer sequences complementary to different ends of a segment of the complementary strands of the target sequence hybridize with their respective strands of the target sequence, and in the presence of polymerase enzymes and nucleoside triphosphates, the primers are extended along the target sequence. The extensions are melted from the target sequence and the process is repeated, this time with the additional copies of the target sequence synthesized in the preceding steps. PCR amplification typically involves repeated cycles of denaturation, hybridization and extension reactions to produce sufficient amounts of the target nucleic acid. The first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

In PCR methods, strand separation is normally achieved by heating the reaction to a sufficiently high temperature for a sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase. Typical heat denaturation involves temperatures ranging from about 80°C to 105° C for times ranging from seconds to minutes. Strand separation, however, can be accomplished by any suitable denaturing method including physical, chemical, or enzymatic means. Strand separation may be induced by a helicase, for example, or an enzyme capable of exhibiting helicase activity.

In addition to PCR reactions, the methods and devices of the present invention are also applicable to a number of other reaction types, e.g., reverse transcription, nick translation, and the like.

PCR may also be used to quantify the amount of transcripts of a particular gene. Typically, quantitative PCR also called real-time PCR is performed on cDNA synthesised from mRNA. Methods of "quantitative" amplification are well known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that may be used to calibrate the PCR reaction. The high density array may then include probes specific to the internal standard for quantification of the amplified nucleic acid.

Thus, in one embodiment, this invention provides a method of detecting of the expression pattern formed by a number of genes. Generally, this method involves providing a high density array containing a multiplicity of probes of one or more particular length(s) that are complementary to subsequences of the mRNA transcribed by the target gene. In one embodiment the high density array may contain every probe of a particular length that is complementary to a particular mRNA. The probes of the high density array are then hybridized with their target nucleic acid alone and then hybridized with a high complexity, high concentration nucleic acid sample that does not contain the targets complementary to the probes. Thus, for example, where the target nucleic acid is an RNA, the probes are first hybridized with their target nucleic acid alone and then hybridized with RNA made from a cDNA library (e.g., reverse transcribed polyA.sup.+ mRNA) where the sense of the hybridized RNA is opposite that of the target nucleic acid (to insure that the

high complexity sample does not contain targets for the probes). Those probes that show a strong hybridization signal with their target and little or no cross-hybridization with the high complexity sample are preferred probes for use in the high density arrays of this invention.

5

The method of measuring the level of expression of a number of genes forming a pattern is, however, not limited to the methods described herein, but includes any quantitative measurement.

#### 10     **Fragmentation**

In addition, amplified sequences may be subjected to other post amplification treatments. For example, in some cases, it may be desirable to fragment the sequence prior to hybridization with an oligonucleotide array, in order to provide segments which are more readily accessible to the probes, which avoid looping and/or hybridization to multiple probes. Fragmentation of the nucleic acids may generally be carried out by physical, chemical or enzymatic methods that are known in the art.

#### 20     **Hybridization**

Following sample preparation, the sample can be subjected to one or more different analysis operations. A variety of analysis operations may generally be performed, including size based analysis using, e.g., microcapillary electrophoresis, and/or sequence based analysis using, e.g., hybridization to an oligonucleotide array.

25     In the latter case, the nucleic acid sample may be probed using an array of oligonucleotide probes. Oligonucleotide arrays generally include a substrate having a large number of positionally distinct oligonucleotide probes attached to the substrate. These arrays may be produced using mechanical or light directed synthesis methods which incorporate a combination of photolithographic methods and solid phase oligonucleotide synthesis methods.

30

#### **Detection**

In the present context high density expression arrays may be used to determine the presence and/or amounts of gene expression products in a sample. Likewise, low



density expression arrays may be used to determine the expression pattern in a sample.

5 While high density expression arrays can be used, other techniques are also contemplated. These include other techniques for assaying for specific mRNA species, including RT-PCR and Northern Blotting, as well as techniques for assaying for particular protein products, such as ELISA, Western blotting, and enzyme assays. Gene expression patterns according to the present invention are determined by measuring any gene product of a particular gene, including mRNA  
10 and protein. A pattern may be for two or more genes.

RNA or protein can be isolated and assayed from a test sample using any techniques known in the art. RNA or protein can for example be isolated from fresh or frozen biopsy, from formalin-fixed tissue.

15 Expression of genes may in general be detected by either detecting mRNA from the cells and/or detecting expression products, such as peptides and proteins.

#### **mRNA detection**

20 The detection of mRNA of the invention may be a tool for determining the developmental stage of a cell type which is defined by its pattern of expression of messenger RNA. For example, in particular stages of cells, high levels of ribosomal RNA are found whereas relatively low levels of other types of messenger RNAs may be found. Where a pattern is shown to be characteristic of a stage, a stage may be defined by that particular pattern of messenger RNA expression. The mRNA  
25 population is a good determinant of developmental stage, will be correlated with other structural features of the cell. In this manner, cells at specific developmental stages will be characterized by the intracellular environment, as well as the extracellular environment. The present invention also allows the combination of definitions based, in part, upon antigens and, in part, upon mRNA expression.  
30

In one embodiment, the two may be combined in a single incubation step. A particular incubation condition may be found which is compatible with both hybridization recognition and non-hybridization recognition molecules. Thus, e.g., an  
35 incubation condition may be selected which allows both specificity of antibody

binding and specificity of nucleic acid hybridization. This allows simultaneous performance of both types of interactions on a single matrix. Again, where developmental mRNA patterns are correlated with structural features, or with probes which are able to hybridize to intracellular mRNA populations, a cell sorter may be used to sort specifically those cells having desired mRNA population patterns.

It is within the general scope of the present invention to provide methods for the detection of mRNA. Such methods often involve sample extraction, PCR amplification, nucleic acid fragmentation and labeling, extension reactions, transcription reactions and the like.

#### **DNA extraction**

DNA extraction may be relevant in case possible mutations in the genes are to be determined in addition to the determination of expression of the genes.

For those embodiments where whole cells, or other tissue samples are being analyzed, it will typically be necessary to extract the nucleic acids from the cells or viruses, prior to continuing with the various sample preparation operations. Accordingly, following sample collection, nucleic acids may be liberated from the collected cells, viral coat, etc., into a crude extract, followed by additional treatments to prepare the sample for subsequent operations, e.g., denaturation of contaminating (DNA binding) proteins, purification, filtration, desalting, and the like.

Liberation of nucleic acids from the sample cells, and denaturation of DNA binding proteins may generally be performed by physical or chemical methods. For example, chemical methods generally employ lysing agents to disrupt the cells and extract the nucleic acids from the cells, followed by treatment of the extract with chaotropic salts such as guanidinium isothiocyanate or urea to denature any contaminating and potentially interfering proteins.

Alternatively, physical methods may be used to extract the nucleic acids and denature DNA binding proteins, such as physical protrusions within microchannels or sharp edged particles piercing cell membranes and extract their contents. Combinations of such structures with piezoelectric elements for agitation can provide suitable shear forces for lysis.

More traditional methods of cell extraction may also be used, e.g., employing a channel with restricted cross-sectional dimension which causes cell lysis when the sample is passed through the channel with sufficient flow pressure. Alternatively, cell extraction and denaturing of contaminating proteins may be carried out by applying an alternating electrical current to the sample. More specifically, the sample of cells is flowed through a microtubular array while an alternating electric current is applied across the fluid flow. Subjecting cells to ultrasonic agitation or forcing cells through microgeometry apertures, thereby subjecting the cells to high shear stress resulting in rupture are also possible extraction methods.

### Filtration

Following extraction, it will often be desirable to separate the nucleic acids from other elements of the crude extract, e.g., denatured proteins, cell membrane particles, salts, and the like. Removal of particulate matter is generally accomplished by filtration, flocculation or the like. Further, where chemical denaturing methods are used, it may be desirable to desalt the sample prior to proceeding to the next step. Desalting of the sample, and isolation of the nucleic acid may generally be carried out in a single step, e.g., by binding the nucleic acids to a solid phase and washing away the contaminating salts or performing gel filtration chromatography on the sample, passing salts through dialysis membranes, and the like. Suitable solid supports for nucleic acid binding include, e.g., diatomaceous earth, silica (i.e., glass wool), or the like. Suitable gel exclusion media, also well known in the art, may also be readily incorporated into the devices of the present invention, and is commercially available from, e.g., Pharmacia and Sigma Chemical.

Alternatively, desalting methods may generally take advantage of the high electrophoretic mobility and negative of DNA compared to other elements. Electrophoretic methods may also be utilized in the purification of nucleic acids from other cell contaminants and debris. Upon application of an appropriate electric field, the nucleic acids present in the sample will migrate toward the positive electrode and become trapped on the capture membrane. Sample impurities remaining free of the membrane are then washed away by applying an appropriate fluid flow. Upon reversal of the voltage, the nucleic acids are released from the membrane in a substantially purer form. Further, coarse filters may also be overlaid on the barriers

to avoid any fouling of the barriers by particulate matter, proteins or nucleic acids, thereby permitting repeated use.

### **Separation of contaminants by chromatography**

5 In a similar aspect, the high electrophoretic mobility of nucleic acids with their negative charges, may be utilized to separate nucleic acids from contaminants by utilizing a short column of a gel or other appropriate matrix or gel which will slow or retard the flow of other contaminants while allowing the faster nucleic acids to pass.

10 This invention provides nucleic acid affinity matrices that bear a large number of different nucleic acid affinity ligands allowing the simultaneous selection and removal of a large number of preselected nucleic acids from the sample. Methods of producing such affinity matrices are also provided. In general the methods involve the steps of a) providing a nucleic acid amplification template array comprising a  
15 surface to which are attached at least 50 oligonucleotides having different nucleic acid sequences, and wherein each different oligonucleotide is localized in a predetermined region of said surface, the density of said oligonucleotides is greater than about 60 different oligonucleotides per 1 cm.<sup>sup.2</sup>, and all of said different oligonucleotides have an identical terminal 3' nucleic acid sequence and an identical  
20 terminal 5' nucleic acid sequence. b) amplifying said multiplicity of oligonucleotides to provide a pool of amplified nucleic acids; and c) attaching the pool of nucleic acids to a solid support.

For example, nucleic acid affinity chromatography is based on the tendency of  
25 complementary, single-stranded nucleic acids to form a double-stranded or duplex structure through complementary base pairing. A nucleic acid (either DNA or RNA) can easily be attached to a solid substrate (matrix) where it acts as an immobilized ligand that interacts with and forms duplexes with complementary nucleic acids present in a solution contacted to the immobilized ligand. Unbound components can  
30 be washed away from the bound complex to either provide a solution lacking the target molecules bound to the affinity column, or to provide the isolated target molecules themselves. The nucleic acids captured in a hybrid duplex can be separated and released from the affinity matrix by denaturation either through heat, adjustment of salt concentration, or the use of a destabilizing agent such as  
35 formamide, TWEEN.TM.-20 denaturing agent, or sodium dodecyl sulfate (SDS).

Affinity columns (matrices) are typically used either to isolate a single nucleic acid typically by providing a single species of affinity ligand. Alternatively, affinity columns bearing a single affinity ligand (e.g. oligo dt columns) have been used to isolate a multiplicity of nucleic acids where the nucleic acids all share a common sequence (e.g. a polyA).

### Affinity matrices

The type of affinity matrix used depends on the purpose of the analysis. For example, where it is desired to analyze mRNA expression levels of particular genes in a complex nucleic acid sample (e.g., total mRNA) it is often desirable to eliminate nucleic acids produced by genes that are constitutively overexpressed and thereby tend to mask gene products expressed at characteristically lower levels. Thus, in one embodiment, the affinity matrix can be used to remove a number of preselected gene products (e.g., actin, GAPDH, etc.). This is accomplished by providing an affinity matrix bearing nucleic acid affinity ligands complementary to the gene products (e.g., mRNAs or nucleic acids derived therefrom) or to subsequences thereof. Hybridization of the nucleic acid sample to the affinity matrix will result in duplex formation between the affinity ligands and their target nucleic acids. Upon elution of the sample from the affinity matrix, the matrix will retain the duplexes nucleic acids leaving a sample depleted of the overexpressed target nucleic acids.

The affinity matrix can also be used to identify unknown mRNAs or cDNAs in a sample. Where the affinity matrix contains nucleic acids complementary to every known gene (e.g., in a cDNA library, DNA reverse transcribed from an mRNA, mRNA used directly or amplified, or polymerized from a DNA template) in a sample, capture of the known nucleic acids by the affinity matrix leaves a sample enriched for those nucleic acid sequences that are unknown. In effect, the affinity matrix is used to perform a subtractive hybridization to isolate unknown nucleic acid sequences. The remaining "unknown" sequences can then be purified and sequenced according to standard methods.

The affinity matrix can also be used to capture (isolate) and thereby purify unknown nucleic acid sequences. For example, an affinity matrix can be prepared that contains nucleic acid (affinity ligands) that are complementary to sequences not

previously identified, or not previously known to be expressed in a particular nucleic acid sample. The sample is then hybridized to the affinity matrix and those sequences that are retained on the affinity matrix are "unknown" nucleic acids. The retained nucleic acids can be eluted from the matrix (e.g. at increased temperature, increased destabilizing agent concentration, or decreased salt) and the nucleic acids can then be sequenced according to standard methods.

Similarly, the affinity matrix can be used to efficiently capture (isolate) a number of known nucleic acid sequences. Again, the matrix is prepared bearing nucleic acids complementary to those nucleic acids it is desired to isolate. The sample is contacted to the matrix under conditions where the complementary nucleic acid sequences hybridize to the affinity ligands in the matrix. The non-hybridized material is washed off the matrix leaving the desired sequences bound. The hybrid duplexes are then denatured providing a pool of the isolated nucleic acids. The different nucleic acids in the pool can be subsequently separated according to standard methods (e.g. gel electrophoresis).

As indicated above the affinity matrices can be used to selectively remove nucleic acids from virtually any sample containing nucleic acids (e.g., in a cDNA library, DNA reverse transcribed from an mRNA, mRNA used directly or amplified, or polymerized from a DNA template, and so forth). The nucleic acids adhering to the column can be removed by washing with a low salt concentration buffer, a buffer containing a destabilizing agent such as formamide, or by elevating the column temperature.

In one particularly preferred embodiment, the affinity matrix can be used in a method to enrich a sample for unknown RNA sequences (e.g. expressed sequence tags (ESTs)). The method involves first providing an affinity matrix bearing a library of oligonucleotide probes specific to known RNA (e.g., EST) sequences. Then, RNA from undifferentiated and/or unactivated cells and RNA from differentiated or activated or pathological (e.g., transformed) or otherwise having a different metabolic state are separately hybridized against the affinity matrices to provide two pools of RNAs lacking the known RNA sequences.

In a preferred embodiment, the affinity matrix is packed into a columnar casing. The sample is then applied to the affinity matrix (e.g. injected onto a column or applied to a column by a pump such as a sampling pump driven by an autosampler). The affinity matrix (e.g. affinity column) bearing the sample is subjected to conditions under which the nucleic acid probes comprising the affinity matrix hybridize specifically with complementary target nucleic acids. Such conditions are accomplished by maintaining appropriate pH, salt and temperature conditions to facilitate hybridization as discussed above.

For a number of applications, it may be desirable to extract and separate messenger RNA from cells, cellular debris, and other contaminants. As such, the device of the present invention may, in some cases, include an mRNA purification chamber or channel. In general, such purification takes advantage of the poly-A tails on mRNA. In particular and as noted above, poly- T oligonucleotides may be immobilized within a chamber or channel of the device to serve as affinity ligands for mRNA. Poly-T oligonucleotides may be immobilized upon a solid support incorporated within the chamber or channel, or alternatively, may be immobilized upon the surface(s) of the chamber or channel itself. Immobilization of oligonucleotides on the surface of the chambers or channels may be carried out by methods described herein including, e.g., oxidation and silanation of the surface followed by standard DMT synthesis of the oligonucleotides.

In operation, the lysed sample is introduced to a high salt solution to increase the ionic strength for hybridization, whereupon the mRNA will hybridize to the immobilized poly-T. The mRNA bound to the immobilized poly-T oligonucleotides is then washed free in a low ionic strength buffer. The poly-T oligonucleotides may be immobilized upon porous surfaces, e.g., porous silicon, zeolites silica xerogels, scintered particles, or other solid supports.

### **Light directed synthesis of oligonucleotide arrays**

The basic strategy for light directed synthesis of oligonucleotide arrays is as follows. The surface of a solid support, modified with photosensitive protecting groups is illuminated through a photolithographic mask, yielding reactive hydroxyl groups in the illuminated regions. A selected nucleotide, typically in the form of a 3'-O-phosphoramidite-activated deoxynucleoside (protected at the 5' hydroxyl with a

photosensitive protecting group), is then presented to the surface and coupling occurs at the sites that were exposed to light. Following capping and oxidation, the substrate is rinsed and the surface is illuminated through a second mask, to expose additional hydroxyl groups for coupling. A second selected nucleotide (e.g., 5'-protected, 3'-O-phosphoramidite-activated deoxynucleoside) is presented to the surface. The selective deprotection and coupling cycles are repeated until the desired set of products is obtained. Since photolithography is used, the process can be readily miniaturized to generate high density arrays of oligonucleotide probes. Furthermore, the sequence of the oligonucleotides at each site is known. See, Pease, et al. Mechanical synthesis methods are similar to the light directed methods except involving mechanical direction of fluids for deprotection and addition in the synthesis steps.

For some embodiments, oligonucleotide arrays may be prepared having all possible probes of a given length. The hybridization pattern of the target sequence on the array may be used to reconstruct the target DNA sequence. Hybridization analysis of large numbers of probes can be used to sequence long stretches of DNA or provide an oligonucleotide array which is specific and complementary to a particular nucleic acid sequence. For example, in particularly preferred aspects, the oligonucleotide array will contain oligonucleotide probes which are complementary to specific target sequences, and individual or multiple mutations of these. Such arrays are particularly useful in the diagnosis of specific disorders which are characterized by the presence of a particular nucleic acid sequence.

Following sample collection and nucleic acid extraction, the nucleic acid portion of the sample is typically subjected to one or more preparative reactions. These preparative reactions include in vitro transcription, labeling, fragmentation, amplification and other reactions. Nucleic acid amplification increases the number of copies of the target nucleic acid sequence of interest. A variety of amplification methods are suitable for use in the methods and device of the present invention, including for example, the polymerase chain reaction method or (PCR), the ligase chain reaction (LCR), self sustained sequence replication (3SR), and nucleic acid based sequence amplification (NASBA).



The latter two amplification methods involve isothermal reactions based on isothermal transcription, which produce both single stranded RNA (ssRNA) and double stranded DNA (dsDNA) as the amplification products in a ratio of approximately 30 or 100 to 1, respectively. As a result, where these latter methods are employed, sequence analysis may be carried out using either type of substrate, i.e., complementary to either DNA or RNA.

Frequently, it is desirable to amplify the nucleic acid sample prior to hybridization. One of skill in the art will appreciate that whatever amplification method is used, if a quantitative result is desired, care must be taken to use a method that maintains or controls for the relative frequencies of the amplified nucleic acids.

#### Labelling of nucleic acids

The nucleic acids in a sample will generally be labeled to facilitate detection in subsequent steps. Labeling may be carried out during the amplification, in vitro transcription or nick translation processes. In particular, amplification, in vitro transcription or nick translation may incorporate a label into the amplified or transcribed sequence, either through the use of labeled primers or the incorporation of labeled dNTPs into the amplified sequence.

Hybridization between the sample nucleic acid and the oligonucleotide probes upon the array is then detected, using, e.g., epifluorescence confocal microscopy. Typically, sample is mixed during hybridization to enhance hybridization of nucleic acids in the sample to nucleic acid probes on the array.

In some cases, hybridized oligonucleotides may be labeled following hybridization. For example, where biotin labeled dNTPs are used in, e.g., amplification or transcription, streptavidin linked reporter groups may be used to label hybridized complexes. Such operations are readily integratable into the systems of the present invention. Alternatively, the nucleic acids in the sample may be labeled following amplification. Post amplification labeling typically involves the covalent attachment of a particular detectable group upon the amplified sequences. Suitable labels or detectable groups include a variety of fluorescent or radioactive labeling groups well known in the art. These labels may also be coupled to the sequences using methods that are well known in the art.

Methods for detection depend upon the label selected. A fluorescent label is preferred because of its extreme sensitivity and simplicity. Standard labeling procedures are used to determine the positions where interactions between a sequence and a reagent take place. For example, if a target sequence is labeled and exposed to a matrix of different probes, only those locations where probes do interact with the target will exhibit any signal. Alternatively, other methods may be used to scan the matrix to determine where interaction takes place. Of course, the spectrum of interactions may be determined in a temporal manner by repeated scans of interactions which occur at each of a multiplicity of conditions. However, instead of testing each individual interaction separately, a multiplicity of sequence interactions may be simultaneously determined on a matrix.

Means of detecting labeled target (sample) nucleic acids hybridized to the probes of the high density array are known to those of skill in the art. Thus, for example, where a colorimetric label is used, simple visualization of the label is sufficient. Where a radioactive labeled probe is used, detection of the radiation (e.g with photographic film or a solid state detector) is sufficient.

In a preferred embodiment, however, the target nucleic acids are labeled with a fluorescent label and the localization of the label on the probe array is accomplished with fluorescent microscopy. The hybridized array is excited with a light source at the excitation wavelength of the particular fluorescent label and the resulting fluorescence at the emission wavelength is detected. In a particularly preferred embodiment, the excitation light source is a laser appropriate for the excitation of the fluorescent label.

The target polynucleotide may be labeled by any of a number of convenient detectable markers. A fluorescent label is preferred because it provides a very strong signal with low background. It is also optically detectable at high resolution and sensitivity through a quick scanning procedure. Other potential labeling moieties include, radioisotopes, chemiluminescent compounds, labeled binding proteins, heavy metal atoms, spectroscopic markers, magnetic labels, and linked enzymes. Another method for labeling may bypass any label of the target sequence. The target may be exposed to the probes, and a double strand hybrid is formed at those positions only. Addition of a double strand specific reagent will detect where

hybridization takes place. An intercalative dye such as ethidium bromide may be used as long as the probes themselves do not fold back on themselves to a significant extent forming hairpin loops. However, the length of the hairpin loops in short oligonucleotide probes would typically be insufficient to form a stable duplex.

5

Suitable chromogens will include molecules and compounds which absorb light in a distinctive range of wavelengths so that a color may be observed, or emit light when irradiated with radiation of a particular wave length or wave length range, e.g., fluorescers. Biliproteins, e.g., phycoerythrin, may also serve as labels.

10

A wide variety of suitable dyes are available, being primarily chosen to provide an intense color with minimal absorption by their surroundings. Illustrative dye types include quinoline dyes, triarylmethane dyes, acridine dyes, alizarine dyes, phthaleins, insect dyes, azo dyes, anthraquinoid dyes, cyanine dyes, phenazathionium dyes, and phenazoxonium dyes. A wide variety of fluorescers may be used either on their own or in conjunction with quencher molecules. Fluorescers of interest fall into a number of categories having certain primary functionalities. These primary functionalities include 1- and 2-aminonaphthalene, p,p'-diaminostilbenes, pyrenes, quaternary phenanthridine salts, 9-aminoacridines, p,p'-diaminobenzophenone imines, anthracenes, oxacarbocyanine, merocyanine, 3-aminoequilenin, perylene, bis-benzoxazole, bis-p-oxazolyl benzene, 1,2-benzophenazin, retinol, bis-3-aminopyridinium salts, hellebrigenin, tetracycline, sterophenol, benzimidzaolylphenylamine, 2-oxo-3-chromen, indole, xanthen, 7-hydroxycoumarin, phenoxazine, salicylate, strophanthidin, porphyrins, triarylmethanes and flavin. Individual fluorescent compounds which have functionalities for linking or which can be modified to incorporate such functionalities include, e.g., dansyl chloride; fluoresceins such as 3,6-dihydroxy-9-phenylxanthhydrol; rhodamineisothiocyanate; N-phenyl 1-amino-8-sulfonatonaphthalene; N-phenyl 2-amino-6-sulfonatonaphthalene; 4-acetamido-4-isothiocyanato-stilbene-2,2'-disulfonic acid; pyrene-3-sulfonic acid; 2-toluidinonaphthalene-6-sulfonate; N-phenyl, N-methyl 2-aminoaphthalene-6-sulfonate; ethidium bromide; stebrine; auromine-0,2-(9'-anthroyl)palmitate; dansyl phosphatidylethanolamine; N,N'-dioctadecyl oxacarbocyanine; N,N'-dihexyl oxacarbocyanine; merocyanine, 4-(3'pyrenyl)butyrate; d-3-aminodesoxy-equilenin; 12

20

25

30

35

12-(9'-anthroyl)stearate; 2-methylantracene; 9-vinylantracene; 2,2'-(vinylene-p-

phenylene)bisbenzoxazole; p-bis[2-(4-methyl-5-phenyl-oxazolyl)]benzene; 6-dimethylamino-1,2-benzophenazin; retinol; bis(3'-aminopyridinium) 1,10-decandiyl diiodide; sulfonaphthylhydrazone of hellibrienin; chlorotetracycline; N-(7-dimethylamino-4-methyl-2-oxo-3-chromenyl)maleimide; N-p-(2-benzimidazolyl)-phenylmaleimide; N-(4-fluoranthyl)maleimide; bis(homovanillic acid); resazarin; 4-chloro-7-nitro-2,1,3-benzooxadiazole; merocyanine 540; resorufin; rose bengal; and 2,4-diphenyl-3(2H)-furanone.

Desirably, fluorescers should absorb light above about 300 nm, preferably about 350 nm, and more preferably above about 400 nm, usually emitting at wavelengths greater than about 10 nm higher than the wavelength of the light absorbed. It should be noted that the absorption and emission characteristics of the bound dye may differ from the unbound dye. Therefore, when referring to the various wavelength ranges and characteristics of the dyes, it is intended to indicate the dyes as employed and not the dye which is unconjugated and characterized in an arbitrary solvent.

Fluorescers are generally preferred because by irradiating a fluorescer with light, one can obtain a plurality of emissions. Thus, a single label can provide for a plurality of measurable events.

Detectable signal may also be provided by chemiluminescent and bioluminescent sources. Chemiluminescent sources include a compound which becomes electronically excited by a chemical reaction and may then emit light which serves as the detectable signal or donates energy to a fluorescent acceptor. A diverse number of families of compounds have been found to provide chemiluminescence under a variety of conditions. One family of compounds is 2,3-dihydro-1,4-phthalazinedione. The most popular compound is luminol, which is the 5-amino compound. Other members of the family include the 5-amino-6,7,8-trimethoxy- and the dimethylamino-calbenz analog. These compounds can be made to luminesce with alkaline hydrogen peroxide or calcium hypochlorite and base. Another family of compounds is the 2,4,5-triphenylimidazoles, with lophine as the common name for the parent product. Chemiluminescent analogs include para-dimethylamino and -methoxy substituents. Chemiluminescence may also be obtained with oxalates, usually oxalyl active esters, e.g., p-nitrophenyl and a peroxide, e.g., hydrogen

peroxide, under basic conditions. Alternatively, luciferins may be used in conjunction with luciferase or lucigenins to provide bioluminescence.

Spin labels are provided by reporter molecules with an unpaired electron spin which can be detected by electron spin resonance (ESR) spectroscopy. Exemplary spin labels include organic free radicals, transitional metal complexes, particularly vanadium, copper, iron, and manganese, and the like. Exemplary spin labels include nitroxide free radicals.

#### 10 **Expression products**

The present invention relates to the classification of cancer in a tissue sample, based on the expression pattern formed by expression products such as mRNA as described above. Furthermore, the invention also relates to determining expression products such as peptides and proteins. The expression products, peptides and proteins, may be detected by any suitable technique known to the person skilled in the art.

In a preferred embodiment the expression products are detected by means of specific antibodies directed to the various expression products, such as immunofluorescent and/or immunohistochemical staining of the tissue.

Immunohistochemical localization of expressed proteins may be carried out by immunostaining of tissue sections from the single tumors to determine which cells expressed the protein encoded by the transcript in question. The transcript levels were used to select a group of proteins supposed to show variation from sample to sample, making possible a rough correlation between level of protein detected and intensity of the transcript on the microarray.

For example tissue sections may be cut from paraffin-embedded tissue blocks, mounted, and deparaffinized by incubation at 80 C° for 10 min, followed by immersion in heated oil at 60 C for 10 min (Estisol 312, Estichem A/S, Denmark) and rehydration. Antigen retrieval is achieved in TEG (TrisEDTA-Glycerol) buffer using microwaves at 900 W. The tissue sections are cooled in the buffer for 15 min before a brief rinse in tap water. Endogenous peroxidase activity is blocked by incubating the sections with 1% H2O2 for 20 min, followed by three rinses in tap

water, 1 min each. The sections are subsequently soaked in PBS buffer for 2 min. The following steps are modified from the descriptions given by Oncogene Science Inc., in the Mouse Immunohistochemistry Detection System, XHCO1 (UniTest, Uniondale, NY, USA). Briefly, the tissue sections are incubated overnight at 4 C with  
5 primary antibody, followed by three rinses in PBS buffer for 5 min each. Afterwards, the sections are incubated with biotinylated secondary antibody for 30 min, rinsed three times with PBS buffer and subsequently incubated with ABC (avidin-biotinylated horseradish peroxidase complex) for 30 min, followed by three rinses in  
10 PBS buffer. Staining is performed by incubation with AEC (3-amino-ethylcarbazole) for 10 min. The tissue sections are counter stained with Mayers hematoxylin, washed in tap water for 5 min. and mounted with glycerol-gelatin. Positive and negative controls may be included in each staining round with all antibodies.

15 In yet another embodiment the expression products may be detected by means of conventional enzyme assays, such as ELISA methods.

Furthermore, the expression products may be detected by means of peptide/protein chips capable of specifically binding the peptides and/or proteins assessed. Thereby an expression pattern may be obtained.

20

### Levels of expression

In the present invention the pattern formed by the expression profiles of genes is used for classifying cancer. The presence and/or amount of a plurality of gene expression products are/is consequently determined. The level of expression of  
25 selected genes is compared between different cancer cells.

The expression level of a particular gene in one cell type may be determined to be at least one fold higher than the expression level in a second cell type. By a one fold higher expression level is meant that the level of expression is doubled, i.e. an  
30 expression level is determined to be of the value 2 in one cell type and thus a one fold higher expression level is 4. By the term two-fold is meant the value is tripled, i.e. an increase from 2 to 6.

35

### Treatment

MSS occurs in 85% of the diagnosed colon cancers, whereas MSI occurs in 15 % of the diagnoses. The choice of treatment in relation to chemotherapeutic agents seems to be related to the microsatellite status of a given colon cancer. MSS cells are affected by treatment using fluorouracil-based drugs (Carethers JM, Hawn MT, Chauhan DP, et al., J Clin Invest 1996;98(1):199-206.; Carethers JM, Chauhan DP, Fink D, et al. Gastroenterology 1999;117(1):123-31; Koi M, Umar A, Chauhan DP, et al. Cancer Res 1994;54(16):4308-12., Hawn MT, Umar A, Carethers JM, et al. Cancer Res 1995;55(17):3721-5). MSI cells are hypersensitive to for example irinotecan and CPT (Hsiang YH, Lihou MG, Liu LF. Cancer Res 1989;49(18):5077-82; Jacob S, Aguado M, Fallik D, Praz F. Cancer Res 2001;61(17):6555-62). Thus, the ability to determine the microsatellite status of a colon cancer can facilitate the selection of a chemotherapeutic agent that will be effective in treatment of that particular colon cancer type.

One aspect of the present invention relates to treatment with anti cancer drugs following the determination of the microsatellite status of a cancer in a sample to be microsatellite stable (MSS) and a prognostic marker. Typically, the MSS status has been determined using the methods of the present invention. A preferred embodiment is treatment using fluorouracil-based drugs, such as 5-fluorouracil, N-methy-N'-nitro-N-nitrosoguanidine or 6-thioguanine. Yet another preferred embodiment is treatment with non-fluorouracil-based anti cancer drugs. Another preferred embodiment of the invention is treatment using anti cancer drugs used in any combination. However; the treatment may also be a combination of treatment using for example fluorouracil-based anti cancer drugs and/or non-fluorouracil-based anti cancer drugs in combination with other sorts of treatment such as surgical intervention, radiation therapy, radiofrequency ablation, immuno therapy, gene therapy.

### Anti cancer drugs

In one embodiment of the method for treatment of an individual comprising the steps of selecting an individual having contracted colon cancer, wherein the microsatellite status is stable and treating the individual with anti cancer drugs that are suitable for the diagnosed nature of the cancer. Fluorouracil-based anti cancer drugs prevent cells from synthesizing DNA and RNA by interfering with the synthesis of nucleic

acids, thus disrupting the growth of the cancer cells. The drugs are therefore also called antimetabolites. The group of fluorouracil-based anti cancer drugs comprises a large number of drugs. One of the drugs currently used in the clinic is 5-fluorouracil that is used to treat several types of cancer including colon cancer. Another preferred drug of the same group is N-methy-N'-nitro-N-nitrosoguanidine, and in particular 6-thioguanine. The drugs belonging to the group of fluorouracil-based anti cancer drugs are not limited to the ones described above, but comprise all fluorouracil-based anti cancer drugs.

- 10 A preferred embodiment for the treatment of MSI cells is the use of non-fluorouracil-based drugs, for example Topoisomerase I-inhibitors such as irinotecan, such as CPT(CPT-II, Camptothecin).

- 15 Another embodiment regards the use of anti cancer drugs that are suitable for cancer treatment but are not fluorouracil-based drugs. The non-fluorouracil based group of drugs is a heterogenous group of drugs that comprises for example antibodies, chemotherapeutic drugs known as antineoplastic drugs, and antidote drugs such as folate. One preferred drug of the group is Cetuximab (Erbix) which is used in the clinic to treat patients with advanced colorectal cancer that has spread to other parts of the body. Erbitux is a monoclonal antibody approved to treat this type of cancer. Another preferred embodiment is a combination treatment of Erbitux to be given intravenously with Irinotecan, another drug approved to fight colorectal cancer. Irinotecan and oxaliplatin are chemotherapy drugs that are given as a treatment for cancer in the colon or rectum and exemplify yet other preferred  
25 embodiments. Irinotecan and oxaliplatin belong to the group of antineoplastics. Leucovorin is another preferred drug belonging to the group of non-fluorouracil-based drugs, where leucovorin is the active form of the B complex vitamin, folate, and is used as an antidote to drugs that decrease levels of folic acid. Some treatments require what is called leucovorin rescue, because the drug used to treat  
30 the cancer has had an adverse effect on folic acid levels.

In a third embodiment is the sequential or combined use of drugs from either of the two groups of drugs.



**Assay**

A further aspect of the invention relates to an assay for classifying cancer in an individual having contracted cancer. At least one marker capable of detecting the microsatellite status in a sample is included in the assay together with at least one marker determining the prognostic marker. The determination of said microsatellite status marker and the determination of the prognostic marker may in the assay be determined sequentially or simultaneously.

In a preferred embodiment the assay comprises at least two markers, one for the microsatellite status and one for the prognostic marker.

The marker (s) is/are preferably specifically detecting a gene as identified herein, in particular the genes of the tables in the examples and as discussed above. However, the marker of microsatellite status may be determined by conventional microsatellite analysis as described elsewhere herein.

As discussed above the marker may be any nucleotide probe, such as a DNA, RNA, PNA, or LNA probe capable of hybridising to mRNA indicative of the expression level. The hybridisation conditions are preferably as described below for probes.

In another embodiment the marker is an antibody capable of specifically binding the expression product in question.

**Detection of expression pattern**

Patterns can be compared manually by a person or by a computer or other machine. An algorithm can be used to detect similarities and differences. The algorithm may score and compare, for example, the genes which are expressed and the genes which are not expressed. Alternatively, the algorithm may look for changes in intensity of expression of a particular gene and score changes in intensity between two samples. Similarities may be determined on the basis of genes which are expressed in both samples and genes which are not expressed in both samples or on the basis of genes whose intensity of expression are numerically similar.

Generally, the detection operation will be performed using a reader device external to the diagnostic device. However, it may be desirable in some cases, to incorporate the data gathering operation into the diagnostic device itself.

- 5 The detection apparatus may be a fluorescence detector, or a spectroscopic detector, or another detector.

Although hybridization is one type of specific interaction which is clearly useful for use in this mapping embodiment, antibody reagents may also be very useful.

10

#### **Data gathering and analysis**

- Gathering data from the various analysis operations, e.g., oligonucleotide and/or microcapillary arrays, will typically be carried out using methods known in the art. For example, the arrays may be scanned using lasers to excite fluorescently labeled targets that have hybridized to regions of probe arrays mentioned above, which can then be imaged using charged coupled devices ("CCDs") for a wide field scanning of the array. Alternatively, another particularly useful method for gathering data from the arrays is through the use of laser confocal microscopy which combines the ease and speed of a readily automated process with high resolution detection.

20

- Following the data gathering operation, the data will typically be reported to a data analysis operation. To facilitate the sample analysis operation, the data obtained by the reader from the device will typically be analyzed using a digital computer. Typically, the computer will be appropriately programmed for receipt and storage of the data from the device, as well as for analysis and reporting of the data gathered, i.e., interpreting fluorescence data to determine the sequence of hybridizing probes, normalization of background and single base mismatch hybridizations, ordering of sequence data in SBH applications, and the like.

25

- 30 It is an object of the present invention to provide a biological sample which may be classified or characterized by analyzing the pattern of specific interactions mentioned above. This may be applicable to a cell or tissue type, to the messenger RNA population expressed by a cell to the genetic content of a cell, or to virtually any sample which can be classified and/or identified by its combination of specific molecular properties.
- 35

**Pharmaceutical composition**

The invention also relates to a pharmaceutical composition for treating the classified cancer, such as colorectal tumors.

5 In one aspect the pharmaceutical composition comprises one or more of the peptides being expression products as defined above. In a preferred embodiment, the peptides are bound to carriers. The peptides may suitably be coupled to a polymer carrier, for example a protein carrier, such as BSA. Such formulations are well-known to the person skilled in the art.

10 The peptides may be suppressor peptides normally lost or decreased in tumor tissue administered in order to stabilise tumors towards a less malignant stage. In another embodiment the peptides are onco-peptides capable of eliciting an immune response towards the tumor cells.

15 In another aspect the pharmaceutical composition comprises genetic material, either genetic material for substitution therapy, or for suppressing therapy as discussed below.

20 In a third aspect the pharmaceutical composition comprises at least one antibody produced as described above.

In the present context the term pharmaceutical composition is used synonymously with the term medicament. The medicament of the invention comprises an effective  
25 amount of one or more of the compounds as defined above, or a composition as defined above in combination with pharmaceutically acceptable additives. Such medicament may suitably be formulated for oral, percutaneous, intramuscular, intravenous, intracranial, intrathecal, intracerebroventricular, intranasal or pulmonic administration. For most indications a localised or substantially localised application is  
30 preferred.

Strategies in formulation development of medicaments and compositions based on the compounds of the present invention generally correspond to formulation strategies for any other protein-based drug product. Potential problems and the guidance  
35 required to overcome these problems are dealt with in several textbooks, e.g.

"Therapeutic Peptides and Protein Formulation. Processing and Delivery Systems",  
Ed. A.K. Banga, Technomic Publishing AG, Basel, 1995.

5       Injectables are usually prepared either as liquid solutions or suspensions, solid  
forms suitable for solution in, or suspension in, liquid prior to injection. The prepara-  
tion may also be emulsified. The active ingredient is often mixed with excipients  
which are pharmaceutically acceptable and compatible with the active ingredient.  
Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol or the  
like, and combinations thereof. In addition, if desired, the preparation may contain  
10       minor amounts of auxiliary substances such as wetting or emulsifying agents, pH  
buffering agents, or which enhance the effectiveness or transportation of the prepa-  
ration.

15       Formulations of the compounds of the invention can be prepared by techniques  
known to the person skilled in the art. The formulations may contain pharmaceuti-  
cally acceptable carriers and excipients including microspheres, liposomes, micro-  
capsules, nanoparticles or the like.

20       The preparation may suitably be administered by injection, optionally at the site,  
where the active ingredient is to exert its effect. Additional formulations which are  
suitable for other modes of administration include suppositories, and, in some  
cases, oral formulations. For suppositories, traditional binders and carriers include  
polyalkylene glycols or triglycerides. Such suppositories may be formed from mix-  
tures containing the active ingredient(s) in the range of from 0.5% to 10%, preferably  
25       1-2%. Oral formulations include such normally employed excipients as, for example,  
pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium  
saccharine, cellulose, magnesium carbonate, and the like. These compositions take  
the form of solutions, suspensions, tablets, pills, capsules, sustained release formu-  
lations or powders and generally contain 10-95% of the active ingredient(s), pref-  
30       erably 25-70%.

35       The preparations are administered in a manner compatible with the dosage formula-  
tion, and in such amount as will be therapeutically effective. The quantity to be ad-  
ministered depends on the subject to be treated, including, e.g. the weight and age  
of the subject, the disease to be treated and the stage of disease. Suitable dosage

ranges are of the order of several hundred  $\mu\text{g}$  active ingredient per administration with a preferred range of from about 0.1  $\mu\text{g}$  to 1000  $\mu\text{g}$ , such as in the range of from about 1  $\mu\text{g}$  to 300  $\mu\text{g}$ , and especially in the range of from about 10  $\mu\text{g}$  to 50  $\mu\text{g}$ . Administration may be performed once or may be followed by subsequent administrations. The dosage will also depend on the route of administration and will vary with the age and weight of the subject to be treated. A preferred dosis would be in the interval 30 mg to 70 mg per 70 kg body weight.

Some of the compounds of the present invention are sufficiently active, but for some of the others, the effect will be enhanced if the preparation further comprises pharmaceutically acceptable additives and/or carriers. Such additives and carriers will be known in the art. In some cases, it will be advantageous to include a compound, which promotes delivery of the active substance to its target.

In many instances, it will be necessary to administrate the formulation multiple times. Administration may be a continuous infusion, such as intraventricular infusion or administration in more doses such as more times a day, daily, more times a week, weekly, etc.

## Therapy

The invention further relates to a method of treating individuals suffering from a classified cancer, in particular for treating a colorectal tumor. In one embodiment of the present invention the tumor cell in question is a microsatellite stable (MSS) tumor cell. In a second embodiment the tumor cell in question is a microsatellite instable (MSI) tumor cell.

In one embodiment the invention relates to a method of substitution therapy, i.e. administration of genetic material generally expressed in MSI cells, but lost or decreased in classified MSS cancer cells. Thus, the invention relates to a method for reducing malignancy of a MSS tumor cell of the Dukes B class, said method comprising

obtaining at least one gene selected from genes being expressed at least one-fold higher in MSI cells than the amount expressed in said MSS tumor cell

introducing said at least one gene into the MSS tumor cell in a manner allowing expression of said gene(s)

5 The at least one gene is preferably selected individually from genes comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM 001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM 004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM 001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM 006275</u>	SFRS6	43	

In a preferred embodiment at least two different genes are introduced into the MSS tumor cell.

10 By the term one-fold is meant that the value is doubled, i.e. an expression level is 2 – a one-fold higher expression level is 4.

By the term two-fold is meant that the value is tripled, i.e. an expression level is 2 – a two-fold higher expression level is 6.

15 In another embodiment the invention relates to a method of substitution therapy, i.e. administration of genetic material generally expressed in MSS cells, but lost or decreased in classified MSI cancer cells. Thus, the invention relates to a method for reducing malignancy of a MSI tumor cell of the Dukes C class, said method comprising

20 obtaining at least one gene selected from genes being expressed at least one-fold higher in MSS cells than the amount expressed in said MSI tumor cell

25 introducing said at least one gene into the MSI tumor cell in a manner allowing expression of said gene(s).

The at least one gene is preferably selected individually from genes comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
-----------	---------	-------------	----------	----

protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 129
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89
hypothetical protein FLJ20618	NM_017903	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	NM_003011.1	SET	103
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104

In a preferred embodiment at least two different genes are introduced into the MSI tumor cell.

- 5 The invention also relates to a method of treating individuals having contracted colon cancer, wherein the microsatellite status has been determined to be stable and the cancer has been determined to be hereditary in nature. Preferably, the determination of microsatellite status and the hereditary nature of the cancer has been determined according to the present invention, analyzing expression patterns as described herein. The method of treating individuals involves introducing at least one gene into the tumor cell, whereby the gene is being expressed. The at least one gene is selected individually from the group of genes listed below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
Homo sapiens mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)	NM_000251	MSH2	136	
Mut L homolog 1	NM_00249.2	MLH1	107	
Homo sapiens PMS1 postmeiotic segregation increased 1 (S. cerevisiae)	NM_000534	PMS1	137	
Homo sapiens PMS2 postmeiotic segregation increased 2 (S. cerevisiae) (PMS2), mRNA	NM_000535	PMS2	138	
Homo sapiens mutS homolog 6 (E. coli)	NM_000179	MSH6	139	

- 15 In a preferred embodiment at least two different genes are introduced into the tumor cell.

- The invention further relates to a method of treating individuals suffering from a classified cancer, in particular for treating a colorectal tumor. In one embodiment of the present invention the tumor cell in question is a microsatellite stable (MSS)

tumor cell. In a second embodiment the tumor cell in question is a microsatellite instable (MSI) tumor cell.

5 In yet another embodiment the invention relates to a method of administration of peptides to the tumor cell in question. The peptide is expressed by at least one gene selected from genes being expressed at least two-fold higher in tumor cells than in the amount expressed in the tumor cells in question. The tumor cell in question is MSI tumor cell or MSI tumor cell.

10 Thus, the invention relates to a method for reducing cell malignancy of a MSS tumor cell of the Dukes B class, said method comprising

15 contacting a a MSS tumor cell with at least one peptide expressed by at least one gene selected from genes being expressed in at least two-fold higher in MSI tumor cells than in the amount expressed in said MSS tumor cell.

The at least one peptide is selected individually from genes as listed below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

20 In a preferred embodiment the invention relates to a method for reducing cell malignancy of a MSI tumor cell of the Dukes C class, said method comprising

25 contacting a MSI tumor cell with at least one peptide expressed by at least one gene selected from genes being expressed in at least two-fold higher in MSS tumor cells than in the amount expressed in said MSI tumor cell.

The at least one peptide is selected individually from genes as listed below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	<u>NM_012408</u> <u>NM_183047</u>	PRKCBP1	57 129	
hepatocellular carcinoma-associated antigen	<u>NM_018487</u>	HCA112	89	



112

hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103

ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104
---------------------------	--------------------	-------	-----

In another aspect the invention relates to a therapy whereby genes generally correlated to disease are inhibited by one or more of the following methods:

- 5 One embodiment of the invention relates to a method for reducing malignancy of a MSS cell of the Dukes B class, said method comprising

obtaining at least one nucleotide probe capable of hybridising with at least one gene of a MSS tumor cell, said at least one gene being selected from genes being  
 10 expressed in an amount at least one-fold lower in MSI tumor cells than the amount expressed in said MSS tumor cell, and

introducing said at least one nucleotide probe into the MSS tumor cell in a manner allowing the probe to hybridise to the at least one gene, thereby inhibiting  
 15 expression of said at least one gene.

The probes are preferably selected from probes capable of hybridising to a nucleotide sequence comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 129	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103	
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104	

20

Another embodiment of the invention relates to a method for reducing malignancy of a MSI tumor cell of the Dukes B class, said method comprising

obtaining at least one nucleotide probe capable of hybridising with at least one gene of a MSI tumor cell, said at least one gene being selected from genes being expressed in an amount at least one-fold lower in MSS tumor cells than the amount expressed in said MSI tumor cell, and

5

introducing said at least one nucleotide probe into the MSI tumor cell in a manner allowing the probe to hybridise to the at least one gene, thereby inhibiting expression of said at least one gene.

10 The probes are preferably selected from probes capable of hybridising to a nucleotide sequence comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

15 These methods are preferably based on anti-sense technology, whereby the hybridisation of said probe to the gene leads to a down-regulation of said gene.

The down-regulation may of course also be based on a probe capable of hybridising to regulatory components of the genes in question, such as promoters.

20 In yet another embodiment the probes consists of the sequences identified above.

The hybridization may be tested in vitro at conditions corresponding to in vivo conditions. Typically, hybridization conditions are of low to moderate stringency. These conditions favour specific interactions between completely complementary sequences, but allow some non-specific interaction between less than perfectly  
25 matched sequences to occur as well. After hybridization, the nucleic acids can be "washed" under moderate or high conditions of stringency to dissociate duplexes that are bound together by some non-specific interaction (the nucleic acids that form these duplexes are thus not completely complementary).

30

As is known in the art, the optimal conditions for washing are determined empirically, often by gradually increasing the stringency. The parameters that can be changed to affect stringency include, primarily, temperature and salt concentration. In general, the lower the salt concentration and the higher the temperature, the higher the stringency. Washing can be initiated at a low temperature (for example, room temperature) using a solution containing a salt concentration that is equivalent to or lower than that of the hybridization solution. Subsequent washing can be carried out using progressively warmer solutions having the same salt concentration. As alternatives, the salt concentration can be lowered and the temperature maintained in the washing step, or the salt concentration can be lowered and the temperature increased. Additional parameters can also be altered. For example, use of a destabilizing agent, such as formamide, alters the stringency conditions.

In reactions where nucleic acids are hybridized, the conditions used to achieve a given level of stringency will vary. There is not one set of conditions, for example, that will allow duplexes to form between all nucleic acids that are 85% identical to one another; hybridization also depends on unique features of each nucleic acid. The length of the sequence, the composition of the sequence (for example, the content of purine-like nucleotides versus the content of pyrimidine-like nucleotides) and the type of nucleic acid (for example, DNA or RNA) affect hybridization. An additional consideration is whether one of the nucleic acids is immobilized (for example, on a filter).

An example of a progression from lower to higher stringency conditions is the following, where the salt content is given as the relative abundance of SSC (a salt solution containing sodium chloride and sodium citrate; 2X SSC is 10-fold more concentrated than 0.2X SSC). Nucleic acids are hybridized at 42°C in 2X SSC/0.1% SDS (sodium dodecylsulfate; a detergent) and then washed in 0.2X SSC/0.1% SDS at room temperature (for conditions of low stringency); 0.2X SSC/0.1% SDS at 42°C (for conditions of moderate stringency); and 0.1X SSC at 68°C (for conditions of high stringency). Washing can be carried out using only one of the conditions given, or each of the conditions can be used (for example, washing for 10-15 minutes each in the order listed above). Any or all of the washes can be repeated. As mentioned above, optimal conditions will vary and can be determined empirically.

35

In another aspect a method of reducing tumorigenicity relates to the use of antibodies against an expression product of a cell from the biological tissue. The antibodies may be produced by any suitable method, such as a method comprising the steps of

- 5 obtaining expression product(s) from at least one gene said gene being expressed as defined below,

immunising a mammal with said expression product(s) obtaining antibodies against the expression product.

10

#### **Use**

The methods described above may be used for producing an assay for classifying cancer in animal tissue.

- 15 Furthermore, the invention relates to the use of a peptide as defined above for preparation of a pharmaceutical composition for the treatment of a classified cancer in animal tissue.

- 20 Furthermore, the invention relates to the use of a gene as defined above for preparation of a pharmaceutical composition for the treatment of a classified cancer in animal tissue.

- 25 Also, the invention relates to the use of a probe as defined above for preparation of a pharmaceutical composition for the treatment of a biological condition in animal tissue.

#### **Gene delivery therapy**

- 30 The genetic material discussed above may be any of the described genes or functional parts thereof. The constructs may be introduced as a single DNA molecule encoding all of the genes, or different DNA molecules having one or more genes. The constructs may be introduced simultaneously or consecutively, each with the same or different markers.

The gene may be linked to the complex as such or protected by any suitable system normally used for transfection such as viral vectors or artificial viral envelope, liposomes or micellas, wherein the system is linked to the complex.

5 Numerous techniques for introducing DNA into eukaryotic cells are known to the skilled artisan. Often this is done by means of vectors, and often in the form of nucleic acid encapsidated by a (frequently virus-like) proteinaceous coat. Gene delivery systems may be applied to a wide range of clinical as well as experimental applications.

10 Vectors containing useful elements such as selectable and/or amplifiable markers, promoter/enhancer elements for expression in mammalian, particularly human, cells, and which may be used to prepare stocks of construct DNAs and for carrying out transfections are well known in the art. Many are commercially available.

15 Various techniques have been developed for modification of target tissue and cells in vivo. A number of virus vectors, discussed below, are known which allow transfection and random integration of the virus into the host. See, for example, Dubensky et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:7529-7533; Kaneda et al., (1989) *Science* 243:375-378; Hiebert et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3594-3598; Hatzoglou et al., (1990) *J. Biol. Chem.* 265:17285-17293; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377-8381. Routes and modes of administering the vector include injection, e.g intravascularly or intramuscularly, inhalation, or other parenteral administration.

25 Advantages of adenovirus vectors for human gene therapy include the fact that recombination is rare, no human malignancies are known to be associated with such viruses, the adenovirus genome is double stranded DNA which can be manipulated to accept foreign genes of up to 7.5 kb in size, and live adenovirus is a safe human vaccine organisms.

30 Another vector which can express the DNA molecule of the present invention, and is useful in gene therapy, particularly in humans, is vaccinia virus, which can be rendered non-replicating (U.S. Pat. Nos. 5,225,336; 5,204,243; 5,155,020; 4,769,330).

35

Based on the concept of viral mimicry, artificial viral envelopes (AVE) are designed based on the structure and composition of a viral membrane, such as HIV-1 or RSV and used to deliver genes into cells in vitro and in vivo. See, for example, U.S. Pat. No. 5,252,348, Schreier H. et al., *J. Mol. Recognit.*, 1995, 8:59-62; Schreier H et al., *J. Biol. Chem.*, 1994, 269:9090-9098; Schreier, H., *Pharm. Acta Helv.* 1994, 68:145-159; Chander, R et al. *Life Sci.*, 1992, 50:481-489, which references are hereby incorporated by reference in their entirety. The envelope is preferably produced in a two-step dialysis procedure where the "naked" envelope is formed initially, followed by unidirectional insertion of the viral surface glycoprotein of interest. This process and the physical characteristics of the resulting AVE are described in detail by Chander et al., (supra). Examples of AVE systems are (a) an AVE containing the HIV-1 surface glycoprotein gp160 (Chander et al., supra; Schreier et al., 1995, supra) or glycosyl phosphatidylinositol (GPI)-linked gp120 (Schreier et al., 1994, supra), respectively, and (b) an AVE containing the respiratory syncytial virus (RSV) attachment (G) and fusion (F) glycoproteins (Stecenko, A. A. et al., *Pharm. Pharmacol. Lett.* 1:127-129 (1992)). Thus, vesicles are constructed which mimic the natural membranes of enveloped viruses in their ability to bind to and deliver materials to cells bearing corresponding surface receptors.

AVEs are used to deliver genes both by intravenous injection and by instillation in the lungs. For example, AVEs are manufactured to mimic RSV, exhibiting the RSV F surface glycoprotein which provides selective entry into epithelial cells. F-AVE are loaded with a plasmid coding for the gene of interest, (or a reporter gene such as CAT not present in mammalian tissue).

The AVE system described herein is physically and chemically essentially identical to the natural virus yet is entirely "artificial", as it is constructed from phospholipids, cholesterol, and recombinant viral surface glycoproteins. Hence, there is no carry-over of viral genetic information and no danger of inadvertent viral infection. Construction of the AVEs in two independent steps allows for bulk production of the plain lipid envelopes which, in a separate second step, can then be marked with the desired viral glycoprotein, also allowing for the preparation of protein cocktail formulations if desired.

Another delivery vehicle for use in the present invention are based on the recent description of attenuated *Shigella* as a DNA delivery system (Sizemore, D. R. et al., Science 270:299-302 (1995), which reference is incorporated by reference in its entirety). This approach exploits the ability of *Shigellae* to enter epithelial cells and escape the phagocytic vacuole as a method for delivering the gene construct into the cytoplasm of the target cell. Invasion with as few as one to five bacteria can result in expression of the foreign plasmid DNA delivered by these bacteria.

A preferred type of mediator of nonviral transfection in vitro and in vivo is cationic (ammonium derivatized) lipids. These positively charged lipids form complexes with negatively charged DNA, resulting in DNA charged neutralization and compaction. The complexes endocytosed upon association with the cell membrane, and the DNA somehow escapes the endosome, gaining access to the cytoplasm. Cationic lipid:DNA complexes appear highly stable under normal conditions. Studies of the cationic lipid DOTAP suggest the complex dissociates when the inner layer of the cell membrane is destabilized and anionic lipids from the inner layer displace DNA from the cationic lipid. Several cationic lipids are available commercially. Two of these, DMRI and DC-cholesterol, have been used in human clinical trials. First generation cationic lipids are less efficient than viral vectors. For delivery to lung, any inflammatory responses accompanying the liposome administration are reduced by changing the delivery mode to aerosol administration which distributes the dose more evenly.

The following are non-limiting examples illustrating the present invention.

### Examples

In the following two complete studies have been performed. The type of experiments performed in Study 1 is described in examples 1 to 6. Study 2 describes an earlier, additional study.

### Study I

101 colorectal tumors were tested for microsatellite instability and their global gene transcription was measured using high-density oligonucleotide microarrays. Unsupervised and supervised classification methods were applied to visualize tumor classes and define sets of genes for classification. Real-time PCR was used to vali-

date the microarray data and to investigate platform independency on an independent set of 47 tumors.

### Patients and biopsy specimens

- 5 A broad spectrum of patients representing different common groups of colorectal cancers were included (Table 15, Fig.1).

**Table 15**

**Summary of clinicopathological and microsatellite features of colon cancer samples.**

Patient group (Danish, Finnish)	N	Median age (range)	Localization in colon		Tumour stage			IHC negative stain	
			right (Danish, Finnish)	left (Danish, Finnish)	N (Danish, Finnish)			N (n tested)	
					0 <sup>a</sup>	II	III	MLH1	MSH2
All cases	118 (44,75)	62.0 (32-87)	44 (7,37)	74 (36,38)	17 (6,11)	36 (14,22)	65 (23,42)	12 (56)	1 (56)
Sporadic microsatellite instable tumors <sup>b</sup>	20 (9,16)	66.8 (44-87)	15 (3,12)	9 (6,4)	-	8 (2,6)	12 (2,10)	6 (11)	0 (11)
Hereditary microsatellite instable tumors <sup>b,c</sup>	17 (4,13)	49.6 (32-75)	9 (2,7)	8 (2,6)	-	10 (2,8)	5 (1,4)	5 (7)	1 (7)
Microsatellite stable tumors	61 (30,36)	61.0 (36-85)	11 (0,11)	55 (30,25)	-	18 (10,8)	48 (20,28)	0 (38)	0 (38)

<sup>a</sup>normal biopsy taken from the resection edge of a tumor

<sup>b</sup>according to microsatellite analysis

<sup>c</sup>according to the Amsterdam criteria

<sup>d</sup>DK is Denmark, SF is Finland)

10

15

20

MSI and MSS tumors both from the right and left colon were used and the MSI tumors represent both sporadic and hereditary cases. Hereditary cases included both missense and truncating germline mutations. All specimens were verified by histology and contained more than 50% tumor tissue. The tumors were resected at 15 different clinics in Denmark and Finland. Patients with stage III disease received adjuvant chemotherapy in addition to curative surgery. None of the patients received pre-operative radiation or chemotherapy. Informed consent was obtained from patients to use their specimens and clinical and pathological data for research purposes and the local ethic committees approved the study.

### Microsatellite-instability analysis



DNA was extracted from microdissected cancer tissue. Control DNA was extracted from blood samples when available, otherwise normal epithelium from the oral resection edge was used. Samples positive for markers BAT25 and BAT26 were scored as microsatellite instable whereas samples positive for only one of these markers were tested for further markers and scored as low-frequency MSI if none of these tested positive. Tumors with low-frequency MSI have similar clinical features as microsatellite stable tumors and were considered as such in this study. For determining the MSI status of the 47-tumors in the test set a pentaplex polymerase chain reaction with five quasimonomorphic mononucleotide repeats was used (uraweera N, Duval A, Reperant M, et al.. Gastroenterology 2002;123(6):1804-11).

### RNA purification

Colon specimens were obtained fresh from surgery and were either immediately snap frozen in liquid nitrogen (Denmark) or transferred to  $-70^{\circ}\text{C}$  freezers with as little delay as possible (Finland). The Finnish samples were stored as dry tissue, whereas the Danish samples were either embedded in OCD-compound or stored in a SDS/guadinium thiocyanate solution. Total RNA was isolated using Trizol (Invitrogen) or GenElute Kits (Sigma) following the manufacturers' instructions.

### Gene expression analysis

Labeling of RNA, hybridization and scanning was performed as described elsewhere (Dyrskjot L, Thykjaer T, Kruhoffer M, et al. Nat Genet 2003;33(1):90-6) Biotin labeled cRNA was prepared from  $10\mu\text{g}$  of total RNA and hybridized to the Human Genome U133A GeneChip (Affymetrix) containing 22,289 probesets. The readings from the quantitative scanning were analyzed by the Affymetrix Software MAS 5.0 and normalized using the quantile normalization procedure implemented in RMA (robust multiarray analysis (Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Nucleic Acids Res 2003;31(4):e15., Bolstad BM, Irizarry RA, Astrand M, Speed TP, Bioinformatics 2003;19(2):185-93)).

### Statistical testing

The 101 samples were divided into four groups according to country of origin and MS status. To study if there was any systematic difference between samples from the two countries, a test statistic  $S_1/S_2$  was used.  $S_1$  is the sum of squared deviations when a gene effect, country effect and microsatellite effect is observed, and  $S_2$  is

the same sum where a gene effect and microsatellite effect is present only. The significance is obtained by calculating  $S_1/S_2$  of data achieved after one hundred random permutations of the country labels. The above test was also performed for each gene separately by considering the number of genes with a test value  $S_1(g)/S_2(g)$  below a given threshold. The same method was used to evaluate differences between the MSI and MSS groups. These calculations were based on the expression level of 5.082 genes with a variance over all tumor samples larger than 0.2.

#### Microsatellite status classifier

The MSI classifier was based on the 5.082 genes defined above. A normal distribution was used with the mean dependent on the gene and the group. For each gene, the variation between the groups and the variation within the groups was calculated to select genes with a high ratio between these. To classify a sample, the sum over the genes of the squared distance from the sample value to the group mean was calculated, standardized by the variance, and assigned the sample to the nearest group. The sample to be classified was excluded when calculating group means and variances.

#### Real-time PCR

The procedures were performed as described ( Birkenkamp-Demtroder K, Christensen LL, Olesen SH, et al. Cancer Res 2002;62(15):4352-63) except that short LNA (Locked Nucleic Acid) enhanced probes from a Human Probe Library (Exiqon™) was used. All samples were normalized to GAPDH as this gene has been reported to be constantly expressed in colorectal cancer samples (Andersen CL, Ledet JL, Orntoft TF, Cancer Res 2004;In press).

#### Classification of new independent test samples based on real-time PCR

To translate the microarray-defined classifier to a PCR-platform, the nine-classifier genes were analyzed by quantitative PCR on a subset of 18 of the 101 tumor samples. The average for each gene and group of the microarray data were multiplied with a constant so that the total average was equal to the average of the corresponding log (Loeb LA., Cancer Res 1991;51(12):3075-9) transformed PCR values. This translation can be made because the normalized PCR values are expected to be proportional to the normalized array values, and on a log scale this becomes an additive difference. The difference is gene specific and is therefore estimated for

each gene separately. Thus, the variation obtained from the microarray data, and used in the classifier, can be used directly on the PCR platform.

### Example 1

#### 5 Hierarchical clustering

The phylogenetic tree resulting from hierarchical clustering is shown with relevant clinical and laboratory data (Fig. 1). The cluster analysis based on 1239 genes with a variation across all samples larger than 0.5 reveals that the samples are mainly separated in accordance to the microsatellite status. On the upper trunk we find two clusters represented mainly by normal biopsies (14/20) and MSS tumors (21/25), respectively. The six tumors among the normal samples were labeled in four different batches, excluding a labeling batch effect.

The lower trunk is divided into a MSI cluster (30/36) and a second MSS cluster (37/37). The MSI cluster contains three morphologically normal tissue specimens. Notably, there is no sign of separation between sporadic and hereditary MSI samples and right sided and left sided tumors are interspersed among each other. The two MSS clusters are either dominated by Danish samples (19/25) or by Finnish samples (26/37), which is likely to be due to differences in sampling and preparation in the two countries.

Based on these observations, we performed a series of statistical tests to evaluate if the observed separation of tumors into MSS and MSI groups as well as into Danish and Finnish groups were significant. Our test value  $S_1/S_2$  was 0.914 between Danish and Finnish and 0.908 between MSI and MSS groups, as compared to minimum values of 0.966 and 0.963, respectively, in 100 permutations, thus demonstrating a very clear separation between the groups. This clear distinction of the groups are caused by many genes even at very strict criteria, i.e. low test statistic  $S_1(g)/S_2(g)$  values (See Table 16).

Table 16

#### Permutation test of groups

Pseudo group	$S_1/S_2$ from data	Smaller values in 100 permutations	Minimum in 100 permutations
DK-SF	0.9146	0	0.9660
MSI-MSS	0.9084	0	0.9631

### Permutation test of genes

Pseudo group		$S_1(j)/S_2(j)$			
		< 0.6	< 0.7	< 0.8	< 0.9
DK-SF	number of genes	22	114	425	1536
	max in 100 permutations	0	1	4	132
MSI-MSS	number of genes	49	151	461	1600
	max in 100 permutations	0	0	13	251

### Example 2

#### Construction of a classifier for microsatellite instability status

To construct the classifier the expression profiles from 101 stage II and III tumors were employed. A maximum likelihood classifier was built in order to select a low number of genes resulting in the best possible separation of the two groups. The performance of the classifier was tested using 1-5082 genes and found to be stable showing 2-5 errors when using 4 to several hundred genes (Fig. 2A). In the final classifier the 9 genes (Table 17) that were most frequently used in the crossvalidation were used which resulted in 3 errors (Fig. 2B). The stability of the classifier was tested by randomly chosen datasets (Fig.2C).

**Table 17. Genes used for the classification of microsatellite status**

Genechip Probe_ID	Gene name	Gene Symbol	Array signal* Microsatellite Stable tumors	Array signal* Microsatellite stable tumors	in-
202072	heterogeneous nuclear ribonucleoprotein L	HNRPL	208 ±73	776 ±340	
203444	metastasis-associated 1-like 1	MTA1L1	45 ±13	104 ±36	
206108	splicing factor, arginine/serine-rich 6	SFRS6	74 ±56	478 ±242	
204533	chemokine (C-X-C motif) ligand 10	CXCL10	111 ±80	315 ±535	
212062	ATPase, class II, type 9a	ATP9a	588 ±222	208 ±114	
218345	hepatocellular carcinoma- associated antigen 112	HCA112	1261 ±603	446 ±271	
222444	hypothetical protein FLJ20618	FLJ20618	776 ±193	338 ±168	
210047	SET translocation (myeloid leukemia-associated)	SET	1351 ±298	478 ±201	
209048	protein kinase C binding protein 1	PRKCBP1	294 ±113	158 ±79	

\* Array signal are median signal intensity values ±standard deviation.

The gene MTA1L1 shown in figures and tables in the examples section is in the context of the present invention identical to MTA2 and has been assigned SEQ ID NO.: 23.

### 5 Example 3

#### Stability of classification

The mean error rate for MSS tumors was 1.62% and for MSI tumors 6.0% for microsatellite instable tumors. More than 50% of the errors were related to three tumors (marked with an asterisk in figure 2B) of which two were wrongly classified in all permutation and one in 94%. The remaining errors were mainly caused by four tumors (marked + in figure 2B) with error rates of 40-47%. In terms of sensitivity and specificity, the classification of MSS tumors could be made with a sensitivity of 98.4% and with a specificity of 94.0% (Table 18). After correction for prevalence of MSS tumors, the positive predictive value of MSS was calculated to be almost 98.5% and the negative predictive value to 91%.

**Table 18**

#### Permutation analysis of the microsatellite instability classifier

	Trainings set	Test set
	Errors in crossvalidation	Test errors
Microsatellite instable tumors	6.48% (n=25, range 0-4)	6.0% (n=9, range 0-3)
Microsatellite stable tumors	1.67% (n=30, range 0-3)	1.62% (n=37, range 0-4)
All	3.85% (n=55, range 0-5)	2.48% (n=46, range 0-4)

### 20 Example 4

#### Construction of a classifier for sporadic versus hereditary microsatellite instable tumors

In order to identify a gene set for identification of hereditary MSI tumors the 19 sporadic MSI samples and 18 HNPCC MSI samples in the data set were subjected to supervised classification as described above (Fig.3A, B). The mismatch repair gene *MLH1* show a general downregulation in sporadic disease whereas *PIWIL1* is lower expressed in hereditary cases (Fig. 3C). Using these two genes only one error occurred: a sporadic MSI tumor was classified as hereditary. Based on t-test 500 permutations were performed to test the significance of these two genes as marker genes and found both genes highly significant with estimated p-values < 0.005.

In addition to identifying MSI tumors, we were also able to classify these as being sporadic or inherited based on the expression of only two genes, *MLH1* and *PIWIL1*. The classification only showed one misclassification out of 37. This single case did not have any family history of CRC but the family is very small and the patient was diagnosed at the age of 32 and may thus represent a missed HNPCC case. A majority of sporadic and about half of hereditary microsatellite instable colorectal cancers are caused by inactivation of *MLH1*. In sporadic tumors this is mostly caused by biallelic promotor hypermethylation whereas somatic mutation or loss of heterozygosity of the wild-type allele is significant mechanisms in hereditary tumors. As a result, the *MLH1* expression level in sporadic disease is strongly compromised whereas one or two alleles of *MLH1* in HNPCC are transcribed although encoding a mutated protein. *PIWIL1* is a member of the human Argonaute family that contain a conserved RNA-binding PAZ domain and may be involved in the development and maintenance of stem cells through the RNA-mediated gene-quelling mechanisms associated with DICER (Yuan Z, Sotsky Kent T, Weber TK., *Oncogene* 2003;22(40):6304-10., Sasaki T, Shiohama A, Minoshima S, Shimizu N., *Genomics* 2003;82(3):323-30. The association of this gene with hereditary MSI tumors is unknown and needs further investigation.

### Example 5

#### Cross platform classification

The gene expression level was measured by real-time PCR of the nine classifier genes in a subset of 18 MSS and MSI samples. Median centered and scaled PCR data gave the same overall picture as clustered array data from the 18 samples (Fig. 4A). As *SET* and *ATP9a* did not work well in the PCR reaction, and only seven of the nine classifier genes (*HNRPL*, *MTA1L1*, *SFR6*, *CXCL10*, *HCA112*, *FLJ20618* and *PRKCBP1*) were used and quantified the transcripts from these genes by real-time PCR in a new independent sample set of 47 tumors containing 35 MSS and 12 MSI. The classifier correctly classified 45 of 47 tumors with only two MSI tumors being classified as MSS (Fig. 4B).

**Example 6****Relation between microsatellite-instability status, stage and survival**

The 9-gene classifier was used to classify 36 patients with Stage II tumors and found that 17 were MSI and 19 MSS. The overall survival was highly significantly related to the classification since 10 of 11 patients that died within five years belonged to the MSS group ( $P=0.0014$ ) (Fig. 5A). Thus, as expected the classifier clearly proved to be a strong predictor of survival in stage II disease and probably can be used to select those MSS patients who may benefit from adjuvant chemotherapy.

Among 65 patients with Stage III tumors receiving adjuvant chemotherapy, 16 were classified as MSI tumors and 49 as MSS tumors. As 6 MSI and 30 MSS patients died within five years of follow-up there was no significant difference in overall survival between these groups ( $P=0.55$ ) (Fig. 5B). A trend was that the patients with MSI tumors showed a poorer short-term survival than those with MSS tumors, contrary to stage II patients. This difference may be in accordance with a recent large study which showed that chemotherapy only benefit the MSS tumor patients ( Ribic CM, Sargent DJ, Moore MJ, et al., N Engl J Med 2003;349(3):247-57).

**Study 2****Background**

Colon cancers microsatellite instability status is a better marker for response to adjuvant chemotherapy with fluorouracil than tumour stage II and III. The majority of hereditary colorectal cancer cases are microsatellite instable. We investigated the possibility of classifying colon tumors based on gene expression in crude biopsies and correlated these to crude survival and investigated if the gene expression profile can also identify hereditary cases from sporadic cases.

**Methods**

Gene transcripts from tumour specimens were quantified using microarray technology. The tumors were clustered using unsupervised and supervised classification algorithms. Sets of genes were defined for classification of microsatellite instability status and sporadic verses hereditary microsatellite instable tumors. Real-time PCR was used to validate microarray data and to investigate platform dependency in a new independent set of 47 colorectal tumors.

## Results

Unsupervised hierarchical clustering revealed that tumors were essentially separated according to microsatellite instability status. Supervised classification of the 97 tumor samples using a maximum likelihood classifier with a crossvalidation loop resulted in tree misclassification as compared to microsatellite analysis using from 106 genes and down to only seven genes. The stability of classification of colon tumors in relation to microsatellite status was tested by permutation analysis. The sensitivity for diagnosis of microsatellite stable tumors exceeded 99% with a specificity exceeding 96%. The positive and negative predictive values exceeded 95% and 98%, respectively. The classifier was demonstrated not to be platform dependent as it could successfully be reproduced by real-time PCR. This was further verified as the classifier also correctly classified 95.7% of a new independent set of 47 colorectal tumors using real-time PCR.

Based on microarray data we identified ten genes that were highly correlated with hereditary disease. Using down to two of these genes 36 of 37 microsatellite unstable tumors could be correctly separated into sporadic and hereditary MSI-H colorectal tumors.

Crude survival according to microsatellite status as determined by the classifier, revealed that stage II colon receiving no adjuvant chemotherapy, that patient displaying microsatellite instability had significantly longer overall survival than patient exhibiting microsatellite stable tumors ( $P=0.0014$ ).

By contrast, the patient with Dukes' C tumors displaying microsatellite instability did not have a significant increase in overall survival as compared to patient exhibiting microsatellite stable tumors ( $P=0.55$ ).

## Conclusion

Colon cancer can be stratified into two molecular distinct groups by quantification of the transcripts of 106 genes or even down to seven genes. The two groups are highly correlated with microsatellite stable (MSS) and microsatellite unstable (MSI) tumors. The 7-gene classifier clearly proved to be a strong predictor of survival in Dukes B and it can be used to select patients who need adjuvant chemotherapy,



namely those classified as MSS. We demonstrate that this classification is also valid when performed by real-time PCR analysis allowing a fast diagnosis in a clinical setting. Finally, sporadic from hereditary cases in tumors exhibiting microsatellite instability can be identified based on gene expression monitoring.

5

### Introduction

Colon is the fourth most frequently diagnosed malignancy and the second most common cause of cancer death in the western world. The standard treatment of colon cancer is advised according to tumor stage. Patient with Dukes' C colon cancer receives a fluorouracil-based adjuvant systemic chemotherapy in addition to surgical resection of the tumor, whereas the treatment for Dukes' B patients is based alone on surgical resection.

There is accumulating evidence that these cancers belong to two distinct molecular types according to genetic alterations. The mutator phenotype featuring tumors with microsatellite instability (MSI) and the suppressor pathway displaying chromosomal instability and microsatellite stable (MSS). MSI has been defined as a change of any length due to either insertions or deletions of repeating units in a microsatellite within a tumor compared to normal tissue and is caused by an underlying defect in the mismatch repair (MMR) system. (Boland et al, CR 1998, 58:5248). The MSI pathway may either be sporadic or hereditary (HNPCC) and whereas the disruption of the MMR system in sporadic MSI tumors is most often caused by somatic methylation of the MLH1 gene promoter more than 90% of HNPCC cancers are caused by germline mutations in MLH1 or MSH2.

25

The MSS pathway to cancer begins with the inactivation of tumor suppressor genes, such as APC/ $\beta$ -catenin genes, followed by activation of oncogenes and inactivation of additional tumor suppressor genes, commonly with a high frequency of allelic losses and cytogenetic abnormalities and abnormal DNA tumor content. Many studies have defined the pathoclinical trait of MSI and MSS tumors and found that MSI positive cancers are most frequently found in the right side of the colon, they tend to be of less differentiated, they tend to be larger in size, are often mucinous and often exhibit extensive infiltration by lymphocytes.

30

Crude survival data suggest that patients with HNPCC have a better prognosis than those with sporadic disease [48,49,50] and studies have also shown that MSI is an independent indicator of good prognosis [35,52,53]. Recently it was shown that MSS benefit from 5-FU treatment/leucovorin treatment (New England J Med, august 2<sup>nd</sup>, 2003) whereas MSI cancer patients gained no advantage in survival.

Gene expression profiling has become an increasingly used method for classification, outcome prediction, prediction of response (for a review see Dyrskjot, expert opinion). Most such studies dealing with colon cancer have dealt with the identification of general tumor markers (Alon U; Levine AJ PNAS (1999); Kitahara O; Tsunoda T., Cancer Res (2001); Notterman DA; Levine AJ, Cancer Res (2001); Yanagawa R; Nakamura (2001); Zou TT; Meltzer SJ, Oncogene (2002), Demtroider CR (2002)), markers for benign adenomas versus adenocarcinomas (Lin YM; Nakamura Y, Oncogene (2002); Williams NS; Becerra C., Clin Cancer Res (2003)), staging (Frederiksen CM; Orntoft TF, J Cancer Res Clin Oncol (2003)), or liver metastasis (Takemasa I; Matsubara K., Biochem Biophys Res Commun (2001); Yanagawa R; Nakamura, Neoplasia (2001); Agrawal D; Yeatman T, J Natl Cancer Inst (2002)). One study has addressed the separation of low-frequency microsatellite instability tumors (MSI-L) from MSI and MSS (PCA) (Mori Y; Meltzer SJ, Cancer Res (2003)). The aim of this study was to build a general applicable and robust classifier based on gene expression to separate MSS from MSI tumors. To achieve such robustness the tumors for this study were collected from 14-16 different clinics, RNA was isolated using different methods and labelled in several batches. Gene expression was measured by DNA microarrays of 101 Danish and Finnish tumors from primary colon cancer patients along with 17 normal biopsies.

## Methods

**Biological material** From the Danish and Finnish CRC tissue banks 101 primary colon cancers and 17 macroscopically normal colon epithelium samples from the oral resection edge were chosen. Only adenocarcinomas from Dukes' stage B and C were included, however, these represented a broad spectrum of tumors in relation to location, heredity, microsatellite instability status, and origin of the patient. All tumors were collected in the period from 1994 to 2002, 68 tumor samples were collected at nine different clinics in Finland and 33 samples were collected at four different clinics in Denmark, 36 were Dukes' B, 67 Dukes' C, 41 were sporadic mi-

crossatellite highly instable (MSI-H) of which were 17 HNPCC, and 59 were sporadic microsatellite stable (MSS) (table 19). None of the patients received pre-operative radiation or chemotherapy.

## 5 Table 19

Summary of clinocopathological and microsatellite features of colon samples

Patient group n (DK,SF)		Median age range	Localization in colon right (DK,SF)   left (DK,SF)		Dukes' Stage			IHC negative stain	
					n (DK,SF)			N (n tested)	
					N <sup>a</sup>	B	C	MLH1	MSH2
All cases	119 (44,75)	62.0	45 (8,37)	74 (36,38)	17 (6,11)	36 (14,22)	66 (20,46)	12 (56)	1 (56)
MSI-H <sup>b</sup>	24 (9,16)	67.0	15 (3,12)	9 (6,4)	-	10 (3,7)	14 (5,9)	6 (11)	0 (11)
HNPCC <sup>c</sup>	17(4,13)	45.0	9 (2,7)	8 (2,6)	-	10 (2,8)	7 (2,5)	6 (8)	1 (8)
MSS	60 (25,35)	63.0	11 (0,11)	49 (25,24)	-	16 (9,7)	44 (16,28)	0 (37)	0 (37)

<sup>a</sup>normal biopsy taken from the resection edge of a tumor

<sup>b</sup>according to microsatellite analysis

<sup>c</sup>all tumors MSI-H<sup>b</sup>

10

**Microsatellite-instability analysis.** From all tumor samples available as paraffin blocks, ten sections were cut at 10µm and stained with haematoxylin. The first and last section was cut at 4 µm and stained with haematoxylin. These two sections were used for the identification of tumor and normal cells from each sample. Regions enriched in tumor cells (more than 90%) were microdissected from these sections and DNA was extracted using a Puregene DNA extraction kit (Gentra Systems, Minneapolis, MN). DNA from blood samples was used as control when available, otherwise normal tissue was microdissected from the tissue sections. The samples were analysed for microsatellite instability according to the NCI guidelines (Boland et al). Samples positive for markers BAT25 and BAT26 were scored as MSI-H. Samples positive for only one of these markers were tested for further markers and scored as MSI-L if none of these tested positive. Since MSI-L has similar clinical features as MSS these samples were considered as MSS in this study. In addition to microsatellite analysis all tumors from which paraffin blocks were available were tested for the presence of MLH1 and MSH2 protein by immunohistochemistry. None of the samples scored MSS were negative for either protein whereas six of the MSI scored samples were positive for both (Table 19).

15

20

25

**RNA purification** Colon specimens were obtained fresh from surgery and were immediately snap frozen in liquid nitrogen either as was, in OCD-compound or in a SDS/guadinium thiocyanate solution. Total RNA was isolated using RNAzol (WAK-Chemie Medical) or spin column technology (Sigma) following the manufactures' instructions.

**Gene expression analysis** These procedures were performed at described in detail elsewhere (Dyrskødt et al). Briefly, ten  $\mu$ g of total RNA was used as starting material for the target preparation as described. First and second strand cDNA synthesis was performed using the SuperScript II System (Invitrogen) according to the manufacturers' instructions except using an oligo-dT primer containing a T7 RNA polymerase promoter site. Labelled aRNA was prepared using the BioArray High Yield RNA Transcript Labelling Kit (Enzo) using Biotin labelled CTP and UTP (Enzo) in the reaction together with unlabeled NTP's. Unincorporated nucleotides were removed using RNeasy columns (Qiagen). Fifteen  $\mu$ g of cRNA was fragmented, loading onto the Affymetrix HG\_U133A probe array cartridge and hybridized for 16h. The arrays were washed and stained in the Affymetrix Fluidics Station and scanned using a confocal laser-scanning microscope (Hewlett Packard GeneArray Scanner G2500A). The readings from the quantitative scanning were analyzed by the Affymetrix Gene Expression Analysis Software (MAS 5.0) and normalized using RMA (robust multi array normalisation, Irizarry et al. 2002) in the statistical application R. Redundant probesets (as defined from Unigene build 168) with high correlation ( $>0.5$ ) over all samples were removed, which reduced the dataset to approximately 14.400 probesets. This dataset was used a source for all further calculations in this manuscript.

#### **Unsupervised agglomerative hierarchical clustering**

For hierarchical cluster analysis 1239 genes with a variation across all samples greater than 0.5 were median-centred to a magnitude of 1. Samples and genes were then clustered using average linkage clustering with a modified Person correlation as similarity metric (Eisen et al., PNAS 95: 14863-14868, 1998). The cluster dendrogram was visualized with TreeView (Eisen).

### Group testing

We make a statistical test where the p-value is evaluated through permutations. For each group and gene we calculate the average and the sum of squared deviations from the average. We then sum these over the genes and the groups:

5

This  
joining DK  
that we

$$S_1 = \sum_{\text{groups}} \sum_{\text{genes}} (X_{ij} - \bar{X}_{gr(i)j})^2$$

expression is calculated for  
with SF and MSI with MSS such  
end up with two groups. The

10

sum of squared deviations is denoted  $S_2$ . As a test statistic we use  $S_1/S_2$ . A small value indicates that there is a real reduction in the deviations when going from 2 to 4 groups and thus the groups have a real significance. To judge if a value is significantly small we use permutations. For each of the four groups left when joining DK and SF we randomly allocate the members to a pseudo DK and pseudo SF in such a way that the number of members in each group are as in the original data.

15

To get an understanding of this separation we performed a test to see if this is caused by few genes or if many genes are involved. For this test we calculated  $S_1 = \sum_{\text{genes}} S_1(\text{gene})$  and similarly with  $S_2 = \sum_{\text{genes}} S_2(\text{gene})$ . For each gene  $j$  we used the test statistic  $S_1(j)/S_2(j)$  (Table 3).

20

### Multidimensional scaling

We carried out multidimensional scaling on median-centered and normalized data using CMD—scale in the statistical application R and visualized in a two-dimensional plot.

25

### Microsatellite status classifier

30

The readings from the quantitative scanning were analyzed by the Affymetrix Gene Expression Analysis Software (MAS 5.0) and normalized using RMA (robust multi array normalisation, Irizarry et al. 2002) in the statistical application R. Redundant probesets (as defined from Unigene build 168) with high correlation ( $>0.5$ ) over all samples were removed, which reduced the dataset to approximately 14.400 probesets.

The microsatellite instability status classifier was based on a dataset of 4.266 genes. These genes result from the removal of genes with a variance over all tumor sam-

ples smaller than 0.2 and genes that separate Danish from Finnish samples with a t-value numerically greater than 2. We used a normal distribution with the mean dependent on the gene and the group (MSI, MSS). For each gene, we calculated the variation between the groups and the variation within the groups to select genes with a high ratio between these. To classify a sample, we calculated the sum over the genes of the squared distance from the sample value to the group mean, standardized by the variance and assigned the sample to the nearest group. The sample to be classified was excluded when calculating group means and variances.

#### Estimation of classifier stability

We validated the performance of the classifier by permutation. One hundred datasets consisting of 30 MSS samples and 25 MSI samples were randomly chosen by permutation for training of the classifier with the remaining samples in each case being assigned to a testset. Averages over the 100 data sets of the number of errors in the cross-validation of the training set and in the test set were used as a measure of the precision of the classifier.

**Real-time PCR (RT-PCR).** The procedures were as described (Birkenkamp-Demtroder) except that we used short LNA (Locked Nucleic Acid) enhanced probes from a Human Probe Library (Exiqon™). In short, cDNA was synthesized from single samples some of which were previously analyzed on GeneChips. Reverse transcription was performed using Superscript II RT (Invitrogen). Real-time PCR analysis was performed on selected genes using the primers (DNA Technology) and probes (Exiqon, DK) described in figure legend X. All samples were normalized to GAPDH as described previously (Birkenkamp-Demtroder et. al. Cancer Res., 62: 4352-4363, 2002).

#### Rebuilding of Classifier based on Real-Time PCR

The 79 tumors samples that were not analysed by real-time PCR were transformed into log ratios using one of the tumor samples as reference and used for training of the classifier. Then 23 samples of which 18 were also analyzed on arrays were equally transformed into log ratios using the same tumor sample as above as reference and tested. The idea behind this translation is that we expect the normalized PCR values to be proportional to the normalized array values, and on a log scale this becomes an additive difference. The difference is gene specific and is therefore

estimated for each gene separately. The variation obtained from the microarray data, and used in the classifier, can be used directly on the PCR platform.

## Results

### 5 Hierarchical clustering

The clinical specimens used in this study were collected in two different countries from 14 different clinics in the period 1994 to 2001. The samples were selected to keep a balanced representation of microsatellite instable (MSI) and microsatellite stable (MSS) tumors from both the right- and left-sided colon. The MSI class was represented both by sporadic MSI and hereditary MSI (HNPCC) tumors. Only Dukes' B and Dukes' C tumor samples were included were selected (table 19). Before any attempt to divide a diverse sample collection into distinct classes analyzed the data for systematic bias that may have been introduced during the experimental procedures. A fast and easy way to discover both true distinct classes as well as systematic biases in the data is to perform a hierarchical clustering.

The phylogenetic tree resulting from hierarchical clustering on 1239 genes (Fig. 6) reveals that the main separating factor is microsatellite status. On the upper trunk we find two clusters represented mainly by normal biopsies (14/21) and MSS tumors (18/25), respectively. The lower trunk is divided into a MSI cluster (30/36) and a second MSS cluster (MSS2-cluster) (34/37). A closer inspection of the two MSS clusters unveil that one is dominated by Danish samples (19/25) and one by Finnish samples (26/37 check). Also, it is worth to notice that the MSI cluster contains a vast majority of Finnish samples (32/36) and that the sporadic MSI samples are interspersed among the hereditary samples. The normal biopsies cluster tight together with a slight tendency to separation according to origin. Tree normal samples cluster within the MSI cluster indicating that resection of these samples may have been to close to the tumor lesion.

Inspection of the gene cluster dendrogram shows that the two groups of MSS tumors are mainly separated by a large cluster of genes being upregulated in the Danish samples (data not shown) indicating that a systematic difference between Danish and Finnish samples.

### Significance of observed groups

Based on these observations, we performed a series of test to evaluate if the observed separation of tumors into MSS and MSI as well as DK and SF are significant. For these tests the tumor samples were grouped into four virtual tumor-groups labelled, i.e. Danish MSI (MSI-DK), Danish MSS (MSS-DK), Finnish MSI (MSI-SF) and Finnish MSS (MSS-SF). Based on 5082 genes with a variance above 0.2, we tested if all four groups are significant or if some of the groups can be joined. We considered the two possibilities of joining DK and SF, and of joining MSI and MSS and made a statistical test where the p-value is evaluated through permutations. In 100 permutations of each group combination our test value  $S_1/S_2$  is considerably smaller than in all permutation (Table 20) demonstrating a very clear separation between DK and SF and between MSI and MSS.

Table 20

Permutation test of groups

Pseudo group	$S_1/S_2$ from data	Smaller values in 100 permutations	Minimum in 100 permutations
DK-SF	0.9072795	0	0.962269
I-S	0.9166195	0	0.9583325

Such a clear distinction between groups may rely on a few highly separating genes or a general difference in the gene expression profile including many genes. For both the DK-SF and MSI-MSS the effect are caused by many genes even at very criteria, i.e. low test statistic  $S_1(j)/S_2(j)$  values (Table 21).

Table 21

Permutation test of genes

Pseudo group		$S_1(j)/S_2(j)$			
		< 0.6	< 0.7	< 0.8	< 0.9
DK-SF	number of genes	36	136	522	1785
	max in 100 permutations	0	0	2	225
MSI-MSS	number of genes	17	103	399	1507
	max in 100 permutations	0	1	8	250

When a property is present that influences a large proportion of the genes this may obscure separation of clinical relevant features in unsupervised clustering. To visual-



ize the effect of such properties, we calculated distances by multidimensional scaling between samples with and without of 816 genes separating DK from SF with a t-value numerically greater than 2 (Fig 7). We see an improved separation of MSI and MSS with Danish and Finnish cases mixed. The MSI-DK samples are not completely separated as they are found both between the MSI-SF and the MSS samples. (These plots are not entirely unsupervised since the groups have been used to remove gene).

### Construction of an MSI-MSS classifier

For the construction of a classifier we used the expression profiles from 97 tumors for which no ambiguity had been identified in relation to microsatellite status. The 816 genes separating DK from SF were excluded, as these would be unreliable for MS classification. We built a maximum likelihood classifier in order to select a minimum of genes giving the largest possible separation of the two groups. We tested the performance of the classifier using 1-1000 genes and found that it was stable showing 3-6 errors when using 4 – 400 genes. Of these 106 genes were especially suited for discrimination of MSS from MSI (table 22).

Table 22

AFFYID	SYMBOL	LOCUSLINK	OMIM	REFSEQ	GENENAME
1405 i at	CCL5	6352	187011	NM_002985	chemokine (C-C motif) ligand 5
200628 s at	WARS	7453	191050	NM_004184	tryptophanyl-tRNA synthetase
200814 at	PSME1	5720	600654	NM_006263	proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)
201641 at	BST2	684	600534	NM_004335	bone marrow stromal cell antigen 2
201649 at	UBE2L6	9246	603890	NM_004223	ubiquitin-conjugating enzyme E2L 6
201674 s at	AKAP1	8165	602449	NM_003488	A kinase (PRKA) anchor protein 1
201762 s at	PSME2	5721	602161	NM_002818	proteasome (prosome, macropain) activator subunit 2 (PA28 beta)
201884 at	CEACAM5	1048	114890	NM_004363	carcinoembryonic antigen-related cell adhesion molecule 5
201910 at	FARP1	10160	602654	NM_005766	FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)
201976 s at	MYO10	4651	601481	NM_012334	myosin X
202072 at	HNRPL	3191	603083	NM_001533	heterogeneous nuclear ribonucleoprotein L
202203 s at	AMFR	267	603243	NM_001144	autocrine motility factor receptor
202262 x at	DDAH2	23564	604744	NM_013974	dimethylarginine dimethylaminohydrolase 2
202510 s at	TNFAIP2	7127	603300	NM_006291	tumor necrosis factor, alpha-induced protein 2
202520 s at	MLH1	4292	120436	NM_000249	mutL homolog 1, colon cancer, non-polyposis type 2 (E. coli)
202589 at	TYMS	7298	188350	NM_001071	thymidylate synthetase

202637 s at	<u>ICAM1</u>	<u>3383</u>	<u>147840</u>	<u>NM_000201</u>	intercellular adhesion molecule 1 (CD human rhinovirus receptor)
202678 at	<u>GTF2A2</u>	<u>2958</u>	<u>600519</u>	<u>NM_004492</u>	general transcription factor IIA, 2, 12l
202762 at	<u>ROCK2</u>	<u>9475</u>	<u>604002</u>	<u>NM_004850</u>	Rho-associated, coiled-coil containin protein kinase 2
203008 x at	<u>APACD</u>	<u>10190</u>		<u>NM_005783</u>	ATP binding protein associated with differentiation
203315 at	<u>NCK2</u>	<u>8440</u>	<u>604930</u>	<u>NM_003581</u>	NCK adaptor protein 2
203335 at	<u>PHYH</u>	<u>5264</u>	<u>602026</u>	<u>NM_006214</u>	phytanoyl-CoA hydroxylase (Refsum ease)
203444 s at	<u>MTA2</u>	<u>9219</u>	<u>603947</u>	<u>NM_004739</u>	metastasis-associated gene family, member 2
203559 s at	<u>ABP1</u>	<u>26</u>	<u>104610</u>	<u>NM_001091</u>	amiloride binding protein 1 (amine oxidase (copper-containing))
203773 x at	<u>BLVRA</u>	<u>644</u>	<u>109750</u>	<u>NM_000712</u>	biliverdin reductase A
203896 s at	<u>PLCB4</u>	<u>5332</u>	<u>600810</u>	<u>NM_000933</u>	phospholipase C, beta 4
203915 at	<u>CXCL9</u>	<u>4283</u>	<u>601704</u>	<u>NM_002416</u>	chemokine (C-X-C motif) ligand 9
204020 at	<u>PURA</u>	<u>5813</u>	<u>600473</u>	<u>NM_005859</u>	purine-rich element binding protein A
204044 at	<u>QPRT</u>	<u>23475</u>	<u>606248</u>	<u>NM_014298</u>	quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylation (carboxylating))
204070 at	<u>RARRES3</u>	<u>5920</u>	<u>605092</u>	<u>NM_004585</u>	retinoic acid receptor responder (tazetene induced) 3
204103 at	<u>CCL4</u>	<u>6351</u>	<u>182284</u>	<u>NM_002984</u>	chemokine (C-C motif) ligand 4
204131 s at	<u>FOXO3A</u>	<u>2309</u>	<u>602681</u>	<u>NM_001455</u>	forkhead box O3A
204326 x at	<u>MT1X</u>	<u>4501</u>	<u>156359</u>	<u>NM_005952</u>	metallothionein 1X
204415 at	<u>G1P3</u>	<u>2537</u>	<u>147572</u>	<u>NM_002038</u>	interferon, alpha-inducible protein (class I FI-6-16)
204533 at	<u>CXCL10</u>	<u>3627</u>	<u>147310</u>	<u>NM_001565</u>	chemokine (C-X-C motif) ligand 10
204745 x at	<u>MT1G</u>	<u>4495</u>	<u>156353</u>	<u>NM_005950</u> , <u>NM_005950</u>	metallothionein 1G
204780 s at	<u>TNFRSF6</u>	<u>355</u>	<u>134637</u>	<u>NM_000043</u> , <u>NM_152877</u> , <u>NM_152876</u> , <u>NM_152875</u> , <u>NM_152872</u> , <u>NM_152873</u> , <u>NM_152871</u>	tumor necrosis factor receptor superfamily, member 6
204858 s at	<u>ECGF1</u>	<u>1890</u>	<u>131222</u>	<u>NM_001953</u>	endothelial cell growth factor 1 (plate derived)
205241 at	<u>SCO2</u>	<u>9997</u>	<u>604272</u>	<u>NM_005138</u>	SCO cytochrome oxidase deficient homolog 2 (yeast)
205242 at	<u>CXCL13</u>	<u>10563</u>	<u>605149</u>	<u>NM_006419</u>	chemokine (C-X-C motif) ligand 13 (E chemoattractant)
205495 s at	<u>GNLY</u>	<u>10578</u>	<u>188855</u>	<u>NM_006433</u> , <u>NM_006433</u>	granulysin
205831 at	<u>CD2</u>	<u>914</u>	<u>186990</u>	<u>NM_001767</u>	CD2 antigen (p50), sheep red blood receptor
206108 s at	<u>SFRS6</u>	<u>6431</u>	<u>601944</u>	<u>NM_006275</u>	splicing factor, arginine/serine-rich 6
206286 s at	<u>TDGF1</u>	<u>6997</u>	<u>187395</u>	<u>NM_003212</u>	teratocarcinoma-derived growth factor
206461 x at	<u>MT1H</u>	<u>4496</u>	<u>156354</u>	<u>NM_005951</u>	metallothionein 1H
206754 s at	<u>CYP2B6</u>	<u>1555</u>	<u>123930</u>	<u>NM_000767</u>	cytochrome P450, family 2, subfamily polypeptide 6
206907 at	<u>TNFSF9</u>	<u>8744</u>	<u>606182</u>	<u>NM_003811</u>	tumor necrosis factor (ligand) superfamily member 9

206918 s at	<u>RBM12</u>	<u>10137</u>	<u>607179</u>	NM_006047, NM_006047	RNA binding motif protein 12
206976 s at	<u>HSPH1</u>	<u>10808</u>		NM_006644	heat shock 105kDa/110kDa protein 1
207320 x at	<u>STAU</u>	<u>6780</u>	<u>601716</u>	NM_004602, NM_004602, NM_017452, NM_017453	staufen, RNA binding protein (Drosophila)
207457 s at	<u>LY6G6D</u>	<u>58530</u>	<u>606038</u>	NM_021246	lymphocyte antigen 6 complex, locus
207993 s at	<u>CHP</u>	<u>11261</u>	<u>606988</u>	NM_007236	calcium binding protein P22
208022 s at	<u>CDC14B</u>	<u>8555</u>	<u>603505</u>	NM_003671, NM_003671, NM_033331	CDC14 cell division cycle 14 homolog (S. cerevisiae)
208156 x at	<u>EPPK1</u>	<u>83481</u>			epiplakin 1
208581 x at	<u>MT1X</u>	<u>4501</u>	<u>156359</u>	NM_005952	metallothionein 1X
208944 at	<u>TGFB2</u>	<u>7048</u>	<u>190182</u>	NM_003242	transforming growth factor, beta receptor II (70/80kDa)
209048 s at	<u>PRKCBP1</u>	<u>23613</u>		NM_012408, NM_012408, NM_183047	protein kinase C binding protein 1
209108 at	<u>TM4SF6</u>	<u>7105</u>	<u>300191</u>	NM_003270	transmembrane 4 superfamily member
209504 s at	<u>PLEKHB1</u>	<u>58473</u>	<u>607651</u>	NM_021200	pleckstrin homology domain containing family B (evectins) member 1
209546 s at	<u>APOL1</u>	<u>8542</u>	<u>603743</u>	NM_003661, NM_003661, NM_145343	apolipoprotein L, 1
210029 at	<u>INDO</u>	<u>3620</u>	<u>147435</u>	NM_002164	indoleamine-pyrrole 2,3 dioxygenase
210103 s at	<u>FOXA2</u>	<u>3170</u>	<u>600288</u>	NM_021784, NM_021784	forkhead box A2
210321 at	<u>GZMH</u>	<u>2999</u>	<u>116831</u>	NM_033423	granzyme H (cathepsin G-like 2, protein h-CCPX)
210538 s at	<u>BIRC3</u>	<u>330</u>	<u>601721</u>	NM_001165, NM_001165	baculoviral IAP repeat-containing 3
211456 x at	<u>AF333388</u>				
212057 at	<u>KIAA0182</u>	<u>23199</u>		XM_050495	KIAA0182 protein
212070 at	<u>GPR56</u>	<u>9289</u>	<u>604110</u>	NM_005682	G protein-coupled receptor 56
212185 x at	<u>MT2A</u>	<u>4502</u>	<u>156360</u>	NM_005953	metallothionein 2A
212229 s at	<u>FBXO21</u>	<u>23014</u>		NM_015002, NM_015002	F-box only protein 21
212336 at	<u>EPB41L1</u>	<u>2036</u>	<u>602879</u>	NM_012156, NM_012156	erythrocyte membrane protein band 4.1 like 1
212341 at	<u>MGC21416</u>	<u>286451</u>		NM_173834	hypothetical protein MGC21416
212349 at	<u>POFUT1</u>	<u>23509</u>	<u>607491</u>	NM_015352, NM_015352	protein O-fucosyltransferase 1
212859 x at	<u>MT1E</u>	<u>4493</u>	<u>156351</u>	NM_175617	metallothionein 1E (functional)
213201 s at	<u>TNNT1</u>	<u>7138</u>	<u>191041</u>	NM_003283, NM_003283, XM_352926	troponin T1, skeletal, slow
213385 at	<u>CHN2</u>	<u>1124</u>	<u>602857</u>	NM_004067	chimerin (chimaerin) 2
213470 s at	<u>HNRPH1</u>	<u>3187</u>	<u>601035</u>	NM_005520	heterogeneous nuclear ribonucleoprotein H1 (H)
213738 s at	<u>ATP5A1</u>	<u>498</u>	<u>164360</u>	NM_004046	ATP synthase, H <sup>+</sup> transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle
213757 at	<u>EIF5A</u>	<u>1984</u>	<u>600187</u>	NM_001970	eukaryotic translation initiation factor

214617 at	PRF1	5551	170280	NM_005041	perforin 1 (pore forming protein)
214924 s at	OIP106	22906	608112	NM_014965	OGT(O-Glc-NAc transferase)-interac protein 106 KDa
215693 x at	DDX27	55661		NM_017895	DEAD (Asp-Glu-Ala-Asp) box polype 27
215780 s at	Hs.382039				
216336 x at	AL031602				
217727 x at	VPS35	55737	606931	NM_018206	vacuolar protein sorting 35 (yeast)
217759 at	TRIM44	54765		NM_017583	tripartite motif-containing 44
217875 s at	TMEPAI	56937	606564	NM_020182, NM_020182, NM_199169, NM_199170	transmembrane, prostate androgen i duced RNA
217917 s at	DNCL2A	83658	607167	NM_014183, NM_014183, NM_177953	dynein, cytoplasmic, light polypeptide
217933 s at	LAP3	51056	170250	NM_015907	leucine aminopeptidase 3
218094 s at	C20orf35	55861		NM_018478, NM_018478	chromosome 20 open reading frame
218237 s at	SLC38A1	81539		NM_030674	solute carrier family 38, member 1
218242 s at	CGI-85	51111		NM_016028, NM_016028	CGI-85 protein
218325 s at	DATF1	11083	604140	NM_022105, NM_022105, NM_080796	death associated transcription factor hepatocellular carcinoma-associated gen 112
218345 at	HCA112	55365		NM_018487	sestrin 1
218346 s at	SESN1	27244	606103	NM_014454	hypothetical protein FLJ20315
218704 at	FLJ20315	54894		NM_017763	hypothetical protein FLJ20647
218802 at	FLJ20647	55013		NM_017918	membrane protein expressed in epit like lung adenocarcinoma
218898 at	CT120	79850		NM_024792	DEAD/H (Asp-Glu-Ala-Asp/His) box p peptide
218943 s at	RIG-I	23586		NM_014314	keratin 23 (histone deacetylase induc
218963 s at	KRT23	25984	606194	NM_015515, NM_015515	UDP-N-acetyl-alpha-D- galactosamine:polypeptide N- acetylgalactosaminyltransferase 6 (G NAc-T6)
219956 at	GALNT6	11226	605148	NM_007210	aryl hydrocarbon receptor nuclear tra locator-like 2
220658 s at	ARNTL2	56938		NM_020183	apobec-1 complementation factor
220951 s at	ACF	29974		NM_014576, NM_014576, NM_138932	hypothetical protein FLJ20232
221516 s at	FLJ20232	54471		NM_019008	apolipoprotein L, 2
221653 x at	APOL2	23780	607252	NM_030882, NM_030882	mitochondrial solute carrier protein
221920 s at	MSCP	51312		NM_016612, NM_016612	hypothetical protein FLJ20618
222244 s at	FLJ20618	55000		NM_017903	

The minimum of three errors was found even using only 7 genes (Table 23).

Table 23. Genes used for the classification of MSS vs MSI tumors

Name	Symbol	Unigene	MSS	MSI
hepatocellular carcinoma-associated antigen 112	HCA112	Hs.12126	1261	653
metastasis-associated 1-like 1	MTA1L1	Hs.173043	45	91
chemokine (C-X-C motif) ligand 10	CXCL10	Hs.2248	104	274
heterogeneous nuclear ribonucleoprotein L	HNRPL	Hs.2730	194	630
hypothetical protein FLJ20618	FLJ20618	Hs.52184	776	388
splicing factor, arginine/serine-rich 6	SFRS6	Hs.6891	74	446
protein kinase C binding protein 1	PRKCBP1	Hs.75871	294	168

## 5 Classification of ambiguous samples

Application of the 7-gene classifier to the four samples showing ambiguity in the microsatellite analyses assigns all four to be microsatellite stable tumor class. Notably, all four showed expression levels of *Tumor Growth Factor  $\beta$  induced protein* (TFGBI), MLH1 and thymidylate synthase (TYMS) that are atypical for MSI tumors. Furthermore, these tumors were all from the left colon. Thus the misclassified tumors are clearly truly MSS or they belong to a yet undefined class of MSI tumors.

### Stability of classification

To estimate the stability of the classifier based on all 97 tumor samples, we generated one hundred new classifiers based on randomly chosen datasets consisting of 30 MSS and 25 MSI samples. In each case the classifiers were tested with the remaining samples. The performance for each set was evaluated and averaged over all 100 training and test sets (Table 24). The mean error rate for MSS tumors was 0.52% and 1.38% for MSI tumors. The seven genes defined above were found to be those genes that were most frequently used in the crossvalidation loop. More than 50% of the errors were related to three tumors of which two were wrongly classified in all permutation and one in 94%. The remaining errors were mainly caused by four tumors with error rates of 40-47% showing that the former three samples are truly assigned contradictory to result from the microsatellite analysis and that four samples could not be assigned with confidence too any of the classes.

Table 24 Performance of the classifier

Trainings set	Test set
Errors in crossvalidation	Test errors

MSI	2.8% (n=25, range 0-6)	1.4% (n=10, range 0-4)
MSS	0.70% (n=30, range 0-3)	0.52% (n=29, range 0-2)
All	1.7% (n=55, range 1-7)	1.9% (n=39, range 0-5)

Table 25

Sensitivity, Specificity, and Predictive Value of Test for MSS based on the eight gene Classifier			
Positive for MSS	True = $(0.9948 \times 29) = 28,8492$	False = $(0.138 \times 10) = 1.38$	
Negative for MSS	False = $(0.0052 \times 29) = 0.1508$	True = $(0.962 \times 10) = 9.62$	
Sensitivity	28.9507/29	=	99.5%
Specificity	9.62/10	=	96.2%
Positive predictive value	28.8492/30.2292	=	95.4%
Negative predictive value	9.62/9.7708	=	98.5%

\*Based on a prevalence for MSS of 85%

5

### Survival classifier

Using the same classification methods described above, we build classifiers for survival based on either all samples or the above defined groups of MSI-H and MSS. As seen in figure 10 a distinction of patient with good prognosis (>5 year survival) from patient with bad prognosis (< 5 years survival) can be achieved with higher precision and using only a fraction of the genes by first separating into MSI-H and MSS groups.

### Construction of a classifier for sporadic versus hereditary microsatellite instable tumors

In order to identify a gene set for identification of hereditary microsatellite instable tumors we applied 19 sporadic microsatellite instable samples and 18 microsatellite instable samples to supervised classification as described above. We found ten genes we high scored for separation of sporadic MSI-H from hereditary MSI-H tumours (Table 26). In crossvalidation we found a minimum number of one error using two genes (Fig 9A) and were used in at least 36 of the 37 crossvalidation loops. The genes were: the mismatch repair gene MLH1 that show a general downregulation in sporadic disease and PIWIL1 that is lower expressed in hereditary cases (Fig 9B). Using these two genes only one error occurred: a sporadic microsatellite instable was classified as hereditary. Based on T-test we performed 500 permutations to test the significance of these two genes for marker genes and found both genes highly significant with p-values < 0.005.

Table 26

AFFYID	SYMBOL	LOCUSL NK	OMIM	REFSEQ	AFFYDESCRIPTION
206194 at	HOXC6	3223	142972	NM_004503	Homeo box C4
214868 at	PIWIL1	9271	605571	NM_004764.2	Piwi (Drosophila)-like 1
					MutL (E. coli) homolog 1 (colon cancer, nonpoly- posis type 2)
202520 s at	MLH1	4292	120436	NM_000249.2	Collapsin response media- tor protein 1
202517 at	CRMP1	1400	602462	NM_001313.2	
205453 at	HOXB2	3212	142967	NM_002145.2	Homeo box B2 (HOXB2)
					Pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyde synthetase)
217791 s at	PYCS/ADH 18A1	5832	138250	NM_002860.2	((PYCS/ADH18A1)
202393 s at	TIEG	7071	601878	NM_005655.1	TGFB inducible early growth response (TIEG)
					Checkpoint with forkhead and ring finger domains (CHFR)
218803 at	CHFR	55743	605209	NM_018223.1	
219877 at	FLJ13842	79698		NM_024645.1	Hypothetical protein FLJ13842 (FLJ13842)
					Phosphoprotein regulated by mitogenic pathways (C8FW)
202241 at	C8FW	10221		NM_025195.2	

5

### Cross platform classification

Real time PCR was applied both to verify the array data and examine if the 7-gene classifier would also perform on this platform. We chose 23 samples of which 18 were also analyzed on arrays. The correlation between the two platforms was high (data not shown). In order to test the performance of classification using PCR data we re-build our classifier with a 79 samples array dataset including only those tumors that were not analyzed with PCR. Two samples were classified in discordance with the microsatellite instability test of which one of them was ambiguously classified by the 7-gene array classifier.

15

### Relation between microsatellite-instability status, stage and survival

Based on the 7-gene classifier, classification of 36 patients with Dukes' B tumors receiving no adjuvant chemotherapy, 18 were classified as MSI tumors and 18 as MSS tumors. The overall survival was highly significantly related to the classification since all nine patients that died within five years of follow-up were belonged to the

20

MSS group ( $P=0.0014$ ) (Fig. 10A). Thus, the 7-gene classifier clearly proved to be a strong predictor of survival in Dukes B and it can be used to select patients who need adjuvant chemotherapy, namely those classified as MSS.

- 5 Among 65 patients with Dukes' C tumors receiving adjuvant chemotherapy, 17 were classified as MSI tumors and as 48 MSS tumors. Of these, 6 MSI and 27 MSS patients died within five years of follow-up meaning no significant difference in overall survival between these groups ( $P=0.55$ ) (Fig. 10B). A trend was that the MSI showed a poorer short-term survival than the MSS, contrary to Dukes B patients.
- 10 This difference can be attributed to the fact that a recent large study has shown that chemotherapy only benefit the MSS tumor patients, thus improving their survival to a level comparable to that which is characteristic of MSI tumor patients.

#### **Clinical application of the discovery**

- 15 In the clinic the 106 or less genes described can be used for predicting outcome of colorectal cancer when examined at the RNA level and also on the protein level as each gene identified is the project is transcribed to RNA that is further translated into protein. The genes can also be used determine which patient should be treated with chemotherapy as only non-microsatellite instable tumors will respond to 5-FU based therapy. Building classifiers can achieve a further stratification of patient with good and bad prognosis after stratification into microsatellite instable and stable tumors.
- 20 The genes used to identify hereditary disease can be used to decide which patient should enter into sequencing analysis of mismatch repair genes.

- 25 The RNA determination can be made in any form using any method that will quantify RNA. The proteins can be measured with any method quantification method that can determine the level of proteins.



## References

- 5 Agrawal D, Chen T, Irby R, Quackenbush J, Chambers AF, Szabo M, Cantor A, Coppola D, Yeatman TJ. Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling. *J Natl Cancer Inst.* 2002 Apr 3;94(7):513-21.
- 10 Birkenkamp-Demtroder K, Christensen LL, Olesen SH, Frederiksen CM, Laiho P, Aaltonen LA, Laurberg S, Sorensen FB, Hagemann R, Orntoft TF. Gene expression in colorectal cancer. *Cancer Res.* 2002 Aug 1;62(15):4352-63.
- 15 Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 1998 Nov 15;58(22):5248-57. Review.
- 20 Chapusot C, Martin L, Bouvier AM, Bonithon-Kopp C, Ecarnot-Laubriet A, Rageot D, Ponnelle T, Laurent Puig P, Faivre J, Piard F. Microsatellite instability and intratumoural heterogeneity in 100 right-sided sporadic colon carcinomas. *Br J Cancer.* 2002 Aug 12;87(4):400-4.
- 25 Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, Wolf H, Orntoft TF. Identifying distinct classes of bladder carcinoma using microarrays. *Nat Genet.* 2003 Jan;33(1):90-6.
- 30 Frederiksen CM, Knudsen S, Laurberg S, Orntoft TF. Classification of Dukes' B and C colorectal cancers using expression arrays. *J Cancer Res Clin Oncol.* 2003 May;129(5):263-71.
- Huang J, Qi R, Quackenbush J, Dauway E, Lazaridis E, Yeatman T. Effects of ischemia on gene expression. *J Surg Res.* 2001 Aug;99(2):222-7.

- Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003 Feb 15;31(4):e15.
- 5 Loukola A, Eklun K, Laiho P, Salovaara R, Kristo P, Jarvinen H, Mecklin JP, Launonen V, Aaltonen LA. Microsatellite marker analysis in screening for hereditary nonpolyposis colorectal cancer (HNPCC). *Cancer Res.* 2001 Jun 1;61(11):4545-9.
- Markowitz S, Hines JD, Lutterbaugh J, Myeroff L, Mackay W, Gordon N, Rustum Y, Luna E, Kleinerman J.
- 10 Mutant K-ras oncogenes in colon cancers Do not predict Patient's chemotherapy response or survival. *Clin Cancer Res.* 1995 Apr;1(4):441-5.
- Mori Y, Selaru FM, Sato F, Yin J, Simms LA, Xu Y, Olaru A, Deacu E, Wang S, Taylor JM, Young J, Leggett B, Jass JR, Abraham JM, Shibata D, Meltzer SJ. The impact of microsatellite instability on the molecular phenotype of colorectal tumors.
- 15 *Cancer Res.* 2003 Aug 1;63(15):4577-82.
- Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, Hamilton SR, Laurent-Puig P, Gryfe R, Shepherd LE, Tu D, Redston M, Gallinger S. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer.
- 20 *N Engl J Med.* 2003 Jul 17;349(3):247-57.

**Claims**

1. A method for classification of cancer in an individual having contracted cancer comprising
  - 5 i) in a sample from the individual having contracted cancer determining the microsatellite status of the tumor and
  - ii) in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount which forms a pattern, determining from said pattern a  
10 prognostic marker, wherein the microsatellite status and the prognostic marker is determined simultaneously or sequentially
  - iii) classifying said cancer from the microsatellite status and the prognostic marker.
- 15 2. The method according to claim 1, wherein the prognostic marker is the hereditary or sporadic nature of said cancer the determination of which comprises the steps of
  - i) in a sample from the individual having contracted cancer, said  
20 sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the hereditary or sporadic nature of said cancer
  - ii) determining the presence and/or amount of said gene expression products forming said pattern,
  - 25 iii) obtaining an indication of the hereditary or sporadic nature of said cancer in the individual based on step ii).
3. The method of claims 1 or 2, wherein the determination of microsatellite  
30 status comprises the steps of
  - i) in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the microsatellite status of said cancer,
  - ii) determining the presence and/or amount of said gene expression products forming said pattern,

iii) obtaining an indication of the microsatellite status of said cancer in the individual based on step ii).

4. The method according to claims 1, 2 or 3, wherein the cancer is colon cancer.

5. The method of any of the preceding claims, wherein a plurality of gene expression products are analysed using solid support, having binding partners (hybridisation partners) for said plurality of gene expression products forming a pattern.

6. The method of any of the preceding claims, wherein a plurality of gene expression products are analysed using binding partners (hybridisation partners) for said plurality of gene expression products forming a pattern.

7. The method of claims 1,2 or 3, wherein at least two of said plurality of gene expression products forming a pattern are used to determine said microsatellite status are selected individually from a group of genes indicative of microsatellite status.

8. The method of claims 1, 2 or 3, wherein at least two of said plurality of gene expression products used to determine the hereditary or sporadic nature of said colon cancer are selected individually from a group of genes indicative for the hereditary or sporadic nature of the cancer.

9. The method of claims 1, 2 or 3, wherein at least two of said plurality of gene expression products forming a pattern used to determine said microsatellite status are selected individually from the group consisting of the genes listed below

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
chemokine (C-C motif) ligand 5	<u>NM_002985</u>	CCL5	1
Tryptophanyl-tRNA synthetase	<u>NM_004184</u>	WARS	2
Proteasome (prosome, macropain) activator subunit 1 (PA28 alpha)	NM_006263	PSME1	3
Bone marrow stromal cell antigen 2	<u>NM_004335</u>	BST2	4
ubiquitin-conjugating enzyme E2L 6	NM_004223	UBE2L6	5

A kinase (PRKA) anchor protein 1	NM_003488	AKAP1	6
Proteasome (prosome, macropain) activator subunit 2 (PA28 beta)	NM_002818	PSME2	7
carcinoembryonic antigen-related cell adhesion molecule 5	NM_004363	CEACAM5	8
FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)	NM_005766	FARP1	9
myosin X	NM_012334	MYO10	10
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11
Autocrine motility factor receptor	NM_001144	AMFR	12
dimethylarginine dimethylaminohydrolase 2	NM_013974	DDAH2	13
tumor necrosis factor, alpha-induced protein 2	NM_006291	TNFAIP2	14
mutL homolog 1, colon cancer, nonpolar type 2 (E. coli)	NM_000249	MLH1	15
thymidylate synthetase	NM_001071	TYMS	16
intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	NM_000201	ICAM1	17
general transcription factor IIA, 2, 12kDa	NM_004492	GTF2A2	18
Rho-associated, coiled-coil containing protein kinase 2	NM_004850	ROCK2	19
ATP binding protein associated with cell differentiation	NM_005783	TXNDC9	20
NCK adaptor protein 2	NM_003581	NCK2	21
phytanoyl-CoA hydroxylase (Refsum disease)	NM_006214	PHYH	22
metastasis-associated gene family, member 2	NM_004739	MTA2	23
amiloride binding protein 1 (amine oxidase (copper-containing))	NM_001091	ABP1	24
Biliverdin reductase A	NM_000712	BLVRA	25
phospholipase C, beta 4	NM_000933	PLCB4	26
chemokine (C-X-C motif) ligand 9	NM_002416	CXCL9	27
purine-rich element binding protein A	NM_005859	PURA	28
quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))	NM_014298	QPRT	29
retinoic acid receptor responder (tazarotene induced) 3	NM_004585	RARRES3	30
chemokine (C-C motif) ligand 4	NM_002984	CCL4	31
forkhead box O3A	NM_001455	FOXO3A	32
interferon, alpha-inducible protein (clone IFI-6-16)	NM_002038	G1P3	34
	NM_022873		123
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35
	NM_005950	MT1G	36
metallothionein 1G	NM_005950		
tumor necrosis factor receptor superfamily, member 6	NM_000043	TNFRSF6	37
	NM_152877		133
	NM_152876		132
	NM_152875		134
	NM_152872		130
	NM_152873		33
	NM_152871		129

	NM_152874		131
endothelial cell growth factor 1 (platelet-derived)	<u>NM_001953</u>	ECGF1	38
SCO cytochrome oxidase deficient homolog 2 (yeast)	<u>NM_005138</u>	SCO2	39
chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant)	<u>NM_006419</u>	CXCL13	40
Granulysin	NM_006433	GNLY	41
CD2 antigen (p50), sheep red blood cell receptor	<u>NM_001767</u>	CD2	42
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43
Teratocarcinoma-derived growth factor 1	<u>NM_003212</u>	TDGF1	44
metallothionein 1H	<u>NM_005951</u>	MT1H	45
cytochrome P450, family 2, subfamily B, polypeptide 6	<u>NM_000767</u>	CYP2B6	46
tumor necrosis factor (ligand) superfamily, member 9	<u>NM_003811</u>	TNFSF9	47
RNA binding motif protein 12	NM_006047	RBM12	48
heat shock 105kDa/110kDa protein 1	<u>NM_006644</u>	HSPH1	49
staufer, RNA binding protein (Drosophila)	NM_004602	STAU	50
	NM_017452		125
	NM_017453		126
lymphocyte antigen 6 complex, locus G6D	<u>NM_021246</u>	LY6G6D	51
calcium binding protein P22	<u>NM_007236</u>	CHP	52
CDC14 cell division cycle 14 homolog B (S. cerevisiae)	<u>NM_003671</u>	CDC14B	53
	NM_033331		115
Epiplakin 1	XM_372063	EPPK1	54
metallothionein 1X	<u>NM_005952</u>	MT1X	55
Transforming growth factor, beta receptor II (70/80kDa)	<u>NM_003242</u>	TGFBR2	56
protein kinase C binding protein 1	NM_012408	PRKCBP1	57
	NM_183047		124
Transmembrane 4 superfamily member 6	<u>NM_003270</u>	TM4SF6	58
pleckstrin homology domain containing, family B (evectins) member 1	<u>NM_021200</u>	PLEKHB1	59
apolipoprotein L, 1	NM_003661	APOL1	60
	NM_145343		120
Indoleamine-pyrrole 2,3 dioxygenase	<u>NM_002164</u>	INDO	61

forkhead box A2	NM_021784	FOXA2	62
granzyme H (cathepsin G-like 2, protein h-CCPX)	<u>NM_033423</u>	GZMH	63
baculoviral IAP repeat-containing 3	NM_001165	BIRC3	64
Homo sapiens metallothionein 1H-like protein		AF333388 (Hs 382039)	135
KIAA0182 protein	<u>NM_014615</u>	KIAA0182	117
G protein-coupled receptor 56	<u>NM_005682</u>	GPR56	65
	<u>NM_201524</u>		116
metallothionein 2A	<u>NM_005953</u>	MT2A	66
F-box only protein 21	NM_015002	FBXO21	67
erythrocyte membrane protein band 4.1-like 1	NM_012156,	EPB41L1	68
	NM_012156		
hypothetical protein MGC21416	<u>NM_173834</u>	MGC21416	69
protein O-fucosyltransferase 1	NM_015352,	POFUT1	70
	NM_015352		
metallothionein 1E (functional)	<u>NM_175617</u>	MT1E	71
tropoin T1, skeletal, slow	NM_003283	TNNT1	72
chimerin (chimaerin) 2	<u>NM_004067</u>	CHN2	73
heterogeneous nuclear ribonucleoprotein H1 (H)	<u>NM_005520</u>	HNRPH1	74
ATP synthase, H <sup>+</sup> transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle	<u>NM_004046</u>	ATP5A1	75
eukaryotic translation initiation factor 5A	<u>NM_001970</u>	EIF5A	76
perforin 1 (pore forming protein)	<u>NM_005041</u>	PRF1	77
OGT(O-GlcNAc transferase)-interacting protein 106 KDa	<u>NM_014965</u>	OIP106	78
DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	<u>NM_017895</u>	DDX27	79
vacuolar protein sorting 35 (yeast)	<u>NM_018206</u>	VPS35	80
tripartite motif-containing 44	<u>NM_017583</u>	TRIM44	81
transmembrane, prostate androgen induced	NM_020182	TMEPAI	82
RNA	NM_199169		127
	NM_199170		128
dynein, cytoplasmic, light polypeptide 2A	NM_014183	DNCL2A	83
	NM_177953		122
leucine aminopeptidase 3	<u>NM_015907</u>	LAP3	84
Chromosome 20 open reading frame 35	NM_018478	C20orf35	85
	NM_033542		118
solute carrier family 38, member 1	<u>NM_030674</u>	SLC38A1	86

CGI-85 protein	NM_016028	CGI-85	87
death associated transcription factor 1	NM_022105,	DATF1	88
	NM_080796		121
hepatocellular carcinoma-associated anti- gen 112	NM_018487	HCA112	89
sestrin 1	NM_014454	SESN1	90
hypothetical protein FLJ20315	NM_017763	FLJ20315	91
hypothetical protein FLJ20647	NM_017918	FLJ20647	92
membrane protein expressed in epithelial- like lung adenocarcinoma	NM_024792	CT120	93
DEAD/H (Asp-Glu-Ala-Asp/His) box poly- peptide	NM_014314	RIG-I	94
keratin 23 (histone deacetylase inducible)	NM_015515,	KRT23	95
UDP-N-acetyl-alpha-D- galactosamine:polypeptide N- acetyl-galactosaminyltransferase 6 (Gal- NAc-T6)	NM_007210	GALNT6	96
aryl hydrocarbon receptor nuclear translo- cator-like 2	NM_020183	ARNTL2	97
apobec-1 complementation factor	NM_014576,	ACF	98
	NM_138932		119
hypothetical protein FLJ20232	NM_019008	FLJ20232	99
apolipoprotein L, 2	NM_030882,	APOL2	100
	NM_145343		120
mitochondrial solute carrier protein	NM_016612	MSCP	101
hypothetical protein FLJ20618	NM_017903	FLJ20618	102
SET translocation (myeloid leukaemia- associated)	NM_003011,	SET	103
	1		
	Xm_030577,		104
ATPase, class II, type 9a	9	ATP9a	

5

10. The method of claims 1, 2 or 3, wherein at least two of said plurality of gene expression products forming a pattern used to determine said microsatellite status are selected individually from the group consisting of the genes listed below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11	
metastasis-associated gene family, member 2	NM_004739	MTA2	23	
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35	
splicing factor, arginine/serine-rich 6	NM_006275	SFRS6	43	



protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89
hypothetical protein FLJ20618	NM_017903	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	NM_003011.1	SET	103
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104

5

11. The method of claims 1, 2 or 3, wherein at least two of said plurality of gene expression products forming a pattern used to determine said microsatellite status are selected individually from the group consisting of the genes listed below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11	
metastasis-associated gene family, member 2	NM_004739	MTA2	23	
chemokine (C-X-C motif) ligand 10	NM_001565	CXCL10	35	
splicing factor, arginine/serine-rich 6	NM_006275	SFRS6	43	
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
hepatocellular carcinoma-associated antigen 112	NM_018487	HCA112	89	
hypothetical protein FLJ20618	NM_017903	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	NM_003011.1	SET	103	
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104	

10

12. The method of claims 1, 2 or 3, wherein

i) at least one of said plurality of gene expression products forming a pattern used to determine said microsatellite status is selected from the group of genes consisting of

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	NM_001533	HNRPL	11	
metastasis-associated gene family, member 2	NM_004739	MTA2	23	

Chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43

and

- ii) at least one of said plurality of gene expression products forming a pattern used to determine said microsatellite status is selected from the group of genes consisting of

5

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103	
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104	

13. The method of claims 1, 2 or 3, wherein

10

- i) at least one of said plurality of gene expression products forming a pattern used to determine said microsatellite status is selected from the group of genes that are down regulated in MSS colon cancers compared to MSI colon cancers consisting of

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35
Splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43

15

and

- ii) at least one of said plurality of gene expression products forming a pattern used to determine said microsatellite status is selected from

the group of genes that are up regulated in MSS colon cancers compared to MSI colon cancers consisting of

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104

- 5 14. The method of claim 13, wherein the difference in the level of the gene expression products forming a pattern is at least one-fold.
15. The method of claim 13, wherein the difference of the level of the gene expression products forming a pattern is at least 1.5 fold.
- 10 16. The method of claim 1, 2 or 3, wherein at least one of said plurality of gene expression products used to determine the hereditary or sporadic nature of said colon cancer are selected individually from the group consisting of the genes as listed below

15

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
Homeo box C6	NM_004503	HOXC6	105
Piwi - like 1	NM_004764.2	PIWIL1	106
Mut L homolog 1	NM_00249.2	MLH1	107
Collapsin response mediator protein 1	NM_001313.2	CRMP1	108
Homeo box B2	NM_002145.2	HOXB2	109
	NM_002860.2	PYCS/ADH18	110
Pyrroline-5-carboxylate synthetase (glutamate gamma-semialdehyd synthetase)		A1	
TGFB inducible early growth response	NM_005655.1	TIEG	111
Checkpoint with forkhead and ring finger domains ??	NM_018223.1	CHFR	112
Hypothetical protein FLJ13842	NM_024645.1	FLJ13842	113
Phosphoprotein regulated by mitogenic pathways	NM_025195.1	C8FW	114

17. The method of claim 1, 2 or 3, wherein at least two of said plurality of gene expression products forming a pattern used to determine said hereditary or sporadic nature of colon cancer are the two genes as listed below

Gene name	Ref seq	Gene symbol	SEQ ID NO.:
Piwi – like 1	NM_004764.2	PIWIL1	106
Mut L homolog 1	NM_00249.2	MLH1	107

18. The method according to claims 1, 2 or 3, wherein the microsatellite status in an individual having contracted colon cancer is microsatellite unstable.

19. The method according to any of the preceding claims, wherein said colon cancer is of Duke's B or Duke's C stage.

20. The method according to any of the preceding claims, wherein said colon cancer is an adenocarcinoma, a carcinoma, a teratoma, a sarcoma, and/or a lymphoma.

21. The method according to any of the preceding claims, wherein the sample is a biopsy of the tissue.

22. The method according to any of the preceding claims, wherein the sample is a cell suspension made from the tissue.

23. The method according to any of the preceding claims, wherein the expression level is determined by determining mRNA of the sample.

24. The method according to any of the preceding claims, wherein the expression level is determined by determining expression products, such as peptides and/or protein in the sample.

25. The method according to any of the preceding claims, wherein the microsatellite status of the colon cancer in an individual has been determined prior to the determination of the presence and/or amount of gene expression products

5

26. The method according to any of the preceding claims, wherein the sporadic or hereditary nature of a colon cancer has been determined prior to the determination of the presence and/or amount of gene expression products.

10

27. A method for classification of cancer in an individual having contracted cancer, wherein the microsatellite status is determined by a method comprising the steps of

15

i) in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the microsatellite status of said cancer,

ii) determining the presence and/or amount of said gene expression products forming said pattern,

20

iii) obtaining an indication of the microsatellite status of said cancer in the individual based on step ii).

28. A method for classification of cancer in an individual having contracted cancer, wherein the hereditary or sporadic nature of the cancer is determined by a method comprising the steps of

25

i) in a sample in a sample from the individual having contracted cancer, said sample comprising a plurality of gene expression products the presence and/or amount of which forms a pattern that is indicative of the hereditary or sporadic nature of said cancer,

30

ii) determining the presence and/or amount of said gene expression products forming said pattern,

iii) obtaining an indication of the hereditary or sporadic nature of said cancer in the individual based on step ii).

35

29. The method according to claim 28, wherein the microsatellite status of said cancer is determined simultaneously or sequentially therewith.

5 30. A method for treatment of an individual comprising the steps of  
i) selecting an individual having contracted a colon cancer, wherein the microsatellite status is stable, determined according to the method of claims 1, 2, 3, 27 or 28  
ii) treating the individual with anti cancer drugs

10 31. The method of treatment according to claim 30, wherein the anti cancer drug is selected from the group of fluorouracil-based drugs.

15 32. The method of treatment according to claim 31, wherein the anti cancer drug is selected from 5-fluorouracil, N-methy-N'-nitro-N-nitrosoguanidine and/or 6-thioguanine.

20 33. The method of treatment according to claim 30, wherein the anti cancer drug is selected from the group of non-fluorouracil based drugs.

34. The method according to claim 33, wherein the anti cancer drug is selected from leucovorin, irinotecan, oxaliplatin, cetuximab.

25 35. A method for treatment of an individual comprising the steps of  
i) selecting an individual having contracted a colon cancer, wherein the microsatellite status is instable, determined according to the method of claims 1, 2, 3, 27 or 34  
ii) treating the individual with anti cancer drugs.

30 36. The method according to claim 35, wherein the anti cancer drug is selected from camptothecin or irinotecan.

35 37. The method according to claim 30 or 35, wherein the microsatellite status has been determined by microsatellite analysis, ELISA, antibody-based histochemical staining, immuno histo chemistry.

38. The method according to claim 30 or 35 wherein the sporadic or hereditary nature of colon cancer has been examined prior to determining the sporadic or hereditary nature of colon cancer by gene expression products forming a pattern.

39. The method according to claim 30 or 35 wherein the sporadic or hereditary nature of colon cancer has been examined by histological examination of the sample.

40. The method according to claim 30 or 35 wherein the sporadic or hereditary nature of colon cancer has been examined by genotyping the sample.

41. A method for reducing malignancy of a cell, said method comprising contacting a tumor cell in question with at least one peptide expressed by at least one gene selected from genes being expressed in an at least two-fold higher in tumor cells than the amount expressed in said tumor cell in question.

42. The method according to claim 41, wherein the at least one peptide is selected individually from genes comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

43. The method according to claim 41, wherein the at least one peptide is selected individually from genes comprising a sequence as identified below

Gene name	Ref seq	Gene	SEQ ID
-----------	---------	------	--------

		symbol	NO.:
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103
ATPase, class II, type 9a	Xm_030577.9	ATP9a	104

44. The method according to claim 41 or 42, wherein the tumor cell is contacted with at least two different peptides.

45. A method for reducing malignancy of a tumor cell in question comprising,

- i) obtaining at least one gene selected from genes being expressed in at least one fold higher in tumor cells than the amount expressed in the tumor cell in question,
- ii) introducing said at least one gene into the tumor cell in question in a manner allowing expression of said gene(s).

46. The method according to claim 45, wherein the at least one gene is selected from genes comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
Heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
Chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

47. The method according to claim 45, wherein the at least one gene is selected from genes comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
Protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 129	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	



SET translocation (myeloid leukaemia-associated)      NM\_003011.1      SET      103

ATPase, class II, type 9a      Xm\_030577.9      ATP9a      104

48. The method according to claim 45, 46 or 47, wherein at least two different genes are introduced into the tumor cell.

49. A method for reducing malignancy of a cell in question, said method comprising

obtaining at least one nucleotide probe capable of hybridising with at least one gene of a tumor cell in question, said at least one gene being selected from genes being expressed in an amount at least one-fold lower in tumor cells than the amount expressed in said tumor cell in question, and

introducing said at least one nucleotide probe into the tumor cell in question in a manner allowing the probe to hybridise to the at least one gene, thereby inhibiting expression of said at least one gene.

50. The method according to claim 49, wherein the nucleotide probe is selected from probes capable of hybridising to a nucleotide sequence comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
protein kinase C binding protein 1	NM_012408 NM_183047	PRKCBP1	57 124	
hepatocellular carcinoma-associated antigen 112	<u>NM_018487</u>	HCA112	89	
hypothetical protein FLJ20618	<u>NM_017903</u>	FLJ20618	102	
SET translocation (myeloid leukaemia-associated)	<u>NM_003011.1</u>	SET	103	
ATPase, class II, type 9a	<u>Xm_030577.9</u>	ATP9a	104	

51. The method according to claim 46, wherein the nucleotide probe is selected from probes capable of hybridising to a nucleotide sequence comprising a sequence as identified below

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
heterogeneous nuclear ribonucleoprotein L	<u>NM_001533</u>	HNRPL	11	
metastasis-associated gene family, member 2	<u>NM_004739</u>	MTA2	23	
chemokine (C-X-C motif) ligand 10	<u>NM_001565</u>	CXCL10	35	
splicing factor, arginine/serine-rich 6	<u>NM_006275</u>	SFRS6	43	

5

52. The method according to claim 49, 50 or 51, wherein at least two different probes are introduced into the tumor cell.

10

53. A method for producing antibodies against an expression product of a cell from a biological tissue, said method comprising the steps of

obtaining expression product(s) from at least one gene said gene being expressed as defined in any of claims 1-29,

15

immunising a mammal with said expression product(s) obtaining antibodies against the expression product.

54. A method for treatment of an individual comprising the steps of

20

i) selecting an individual having contracted a colon cancer, wherein the microsatellite status is stable, determined according to the method of claims 1, 2, 3, 27 or 28 and wherein the hereditary nature of said cancer has been determined according to the method of claims 1, 2 or 3

25

ii) introducing at least one gene into the tumor cell in a manner allowing expression of said gene(s).

55. The method according to claim 54, wherein the at least one gene is selected from MSH2, MLH1, PMS1, PMS2 or MSH6.

30

Gene name	Ref seq	Gene symbol	SEQ NO.:	ID
	NM_000251	MSH2	136	
Homo sapiens mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)				
Mut L homolog 1	NM_00249.2	MLH1	107	
Homo sapiens PMS1 postmeiotic segregation increased 1 (S. cerevisiae)	NM_000534	PMS1	137	
Homo sapiens PMS2 postmeiotic segregation increased 2 (S. cerevisiae) (PMS2), mRNA	NM_000535	PMS2	138	
Homo sapiens mutS homolog 6 (E. coli)	NM_000179	MSH6	139	

56. The method according to claim 54 or 55, wherein at least two different genes are introduced.

5 57. Pharmaceutical composition for the treatment of a classified cancer comprising at least one antibody as defined in claim 53.

10 58. Pharmaceutical composition for the treatment of a classified cancer comprising at least one polypeptide as defined in any of the claims 41-44.

15 59. Pharmaceutical composition for the treatment of a classified cancer comprising at least one nucleic acid and/or probe as defined in any of the claims 45-52.

20 60. The use of a method as defined in any of claims 1- 37 for producing an assay for classifying cancer in animal tissue.

25 61. The use of a peptide as defined in any of claims 41-44 for preparation of a pharmaceutical composition for the treatment of a cancer in animal tissue.

62. The use of a gene as defined in any of claims 45-52 for preparation of a pharmaceutical composition for the treatment of cancer in animal tissue.

25

63. The use of a probe as defined in any of claims 49-52 for preparation of a pharmaceutical composition for the treatment of cancer in animal tissue.

5 64. An assay for classification of cancer in an individual having contracted cancer, comprising

at least one marker capable of determining the microsatellite status in a sample and

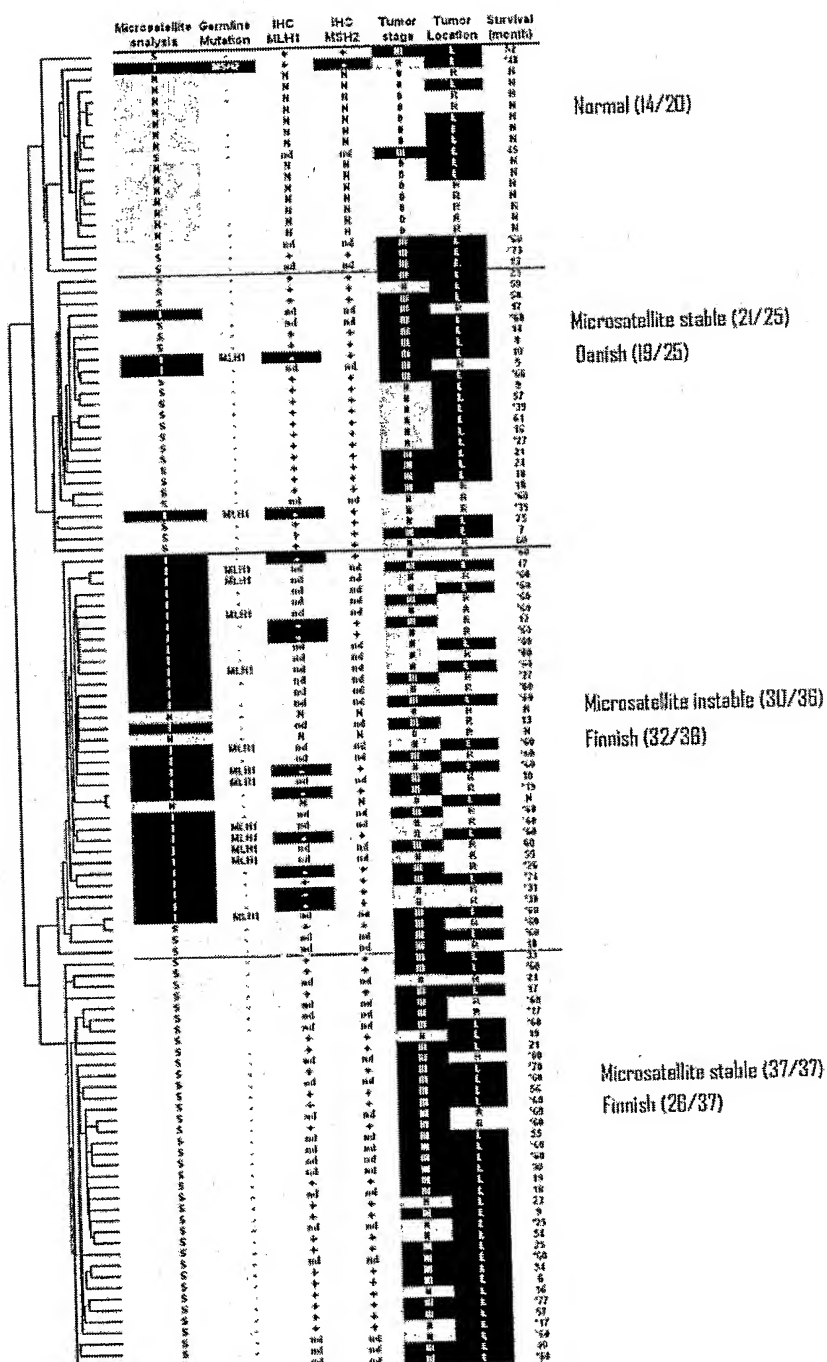
10 at least one marker in a sample determining the prognostic marker, wherein the microsatellite status and the prognostic marker is determined simultaneously or sequentially.

15 65. The assay according to claim 64, wherein the marker is a nucleotide probe.

66. The assay according to claim 64, wherein the marker is an antibody.

20 67. The assay according to claim 64, wherein the genes are as defined in any of claims 9-13 or 16-17.

Fig. 1/10



26 NOV. 2004

PVS

Fig. 2/10

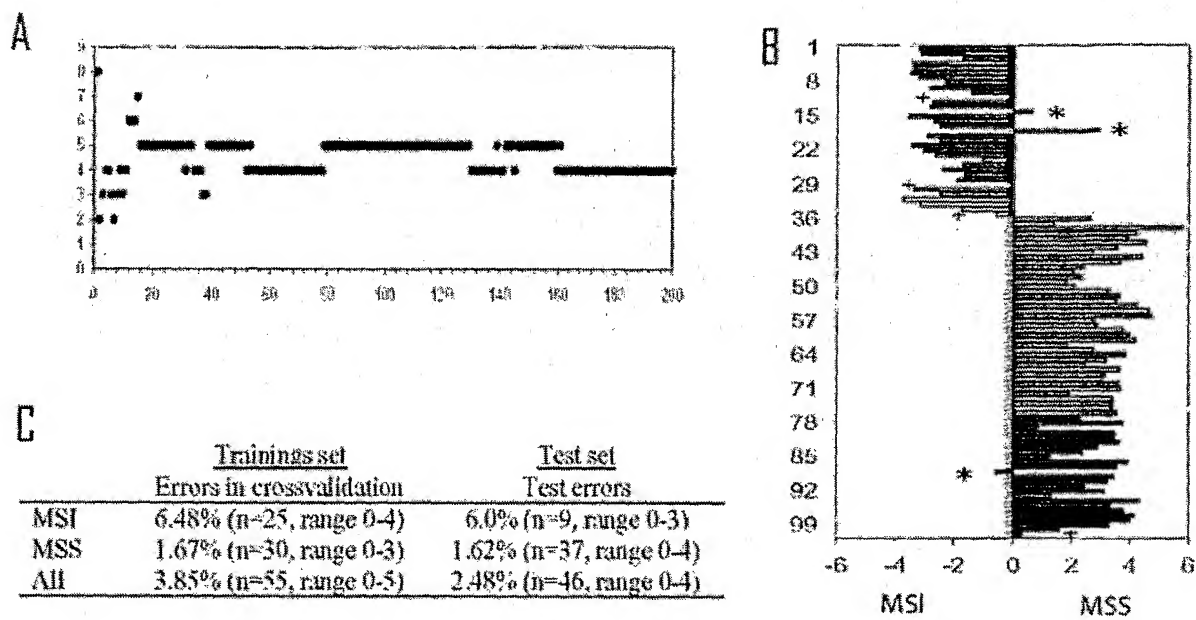
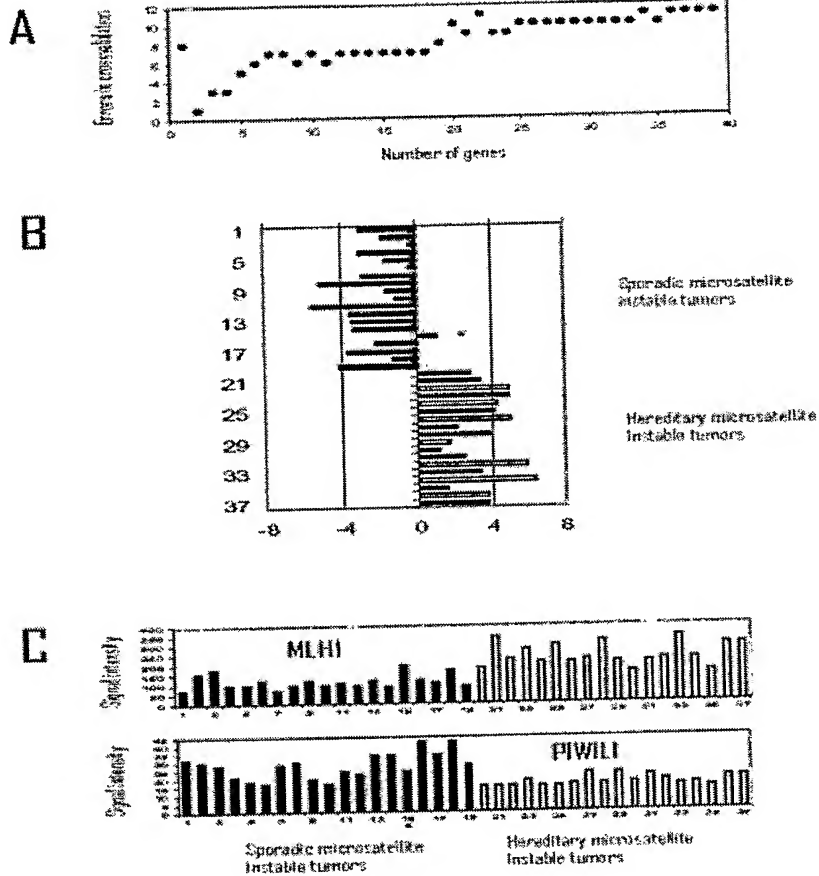


Fig. 3/10



26 NOV. 2004

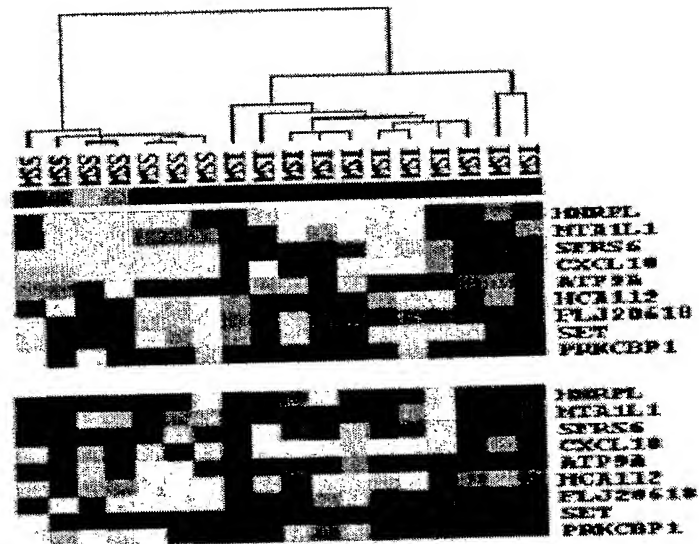
PVS

Fig. 4/10

A

Microarray

PCR



B

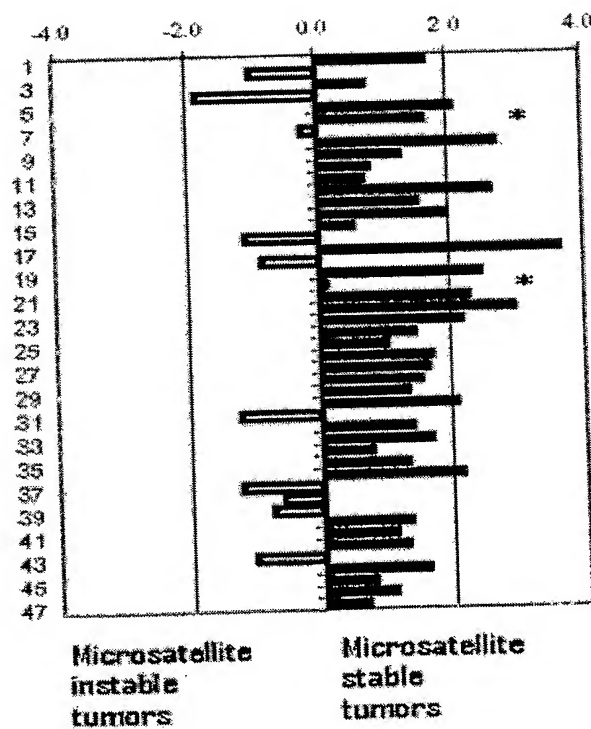
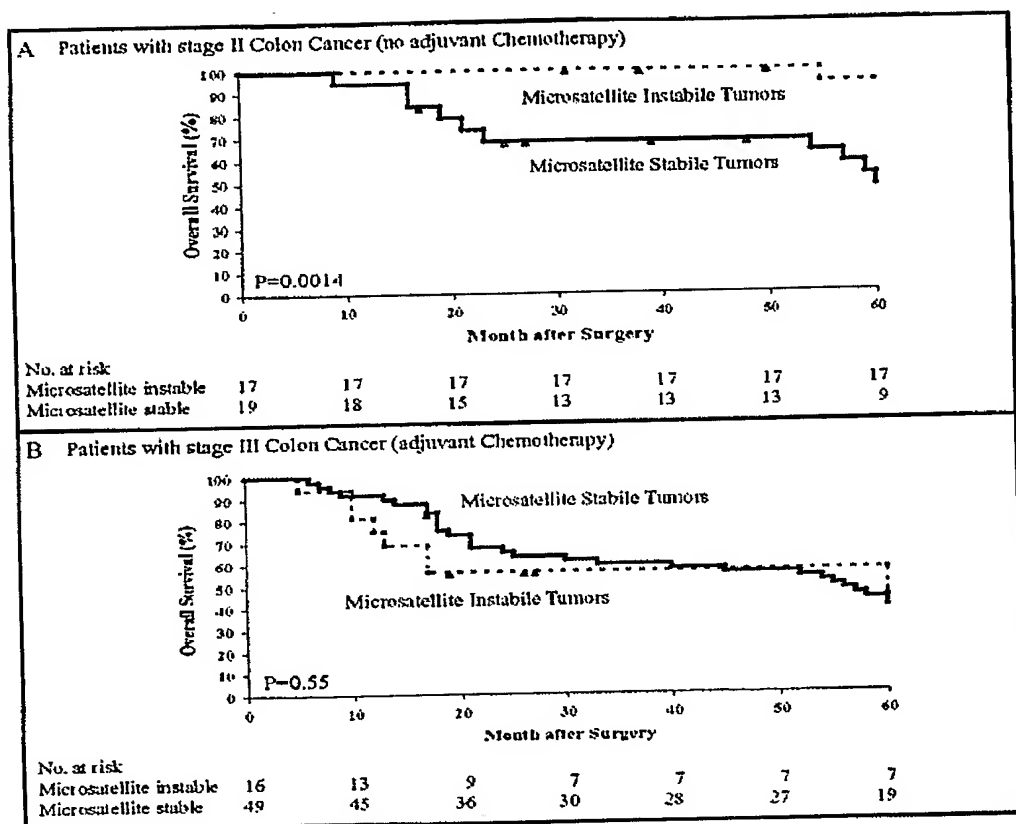




Fig. 5/10



26 NOV. 2004

PVS

Fig. 6/10

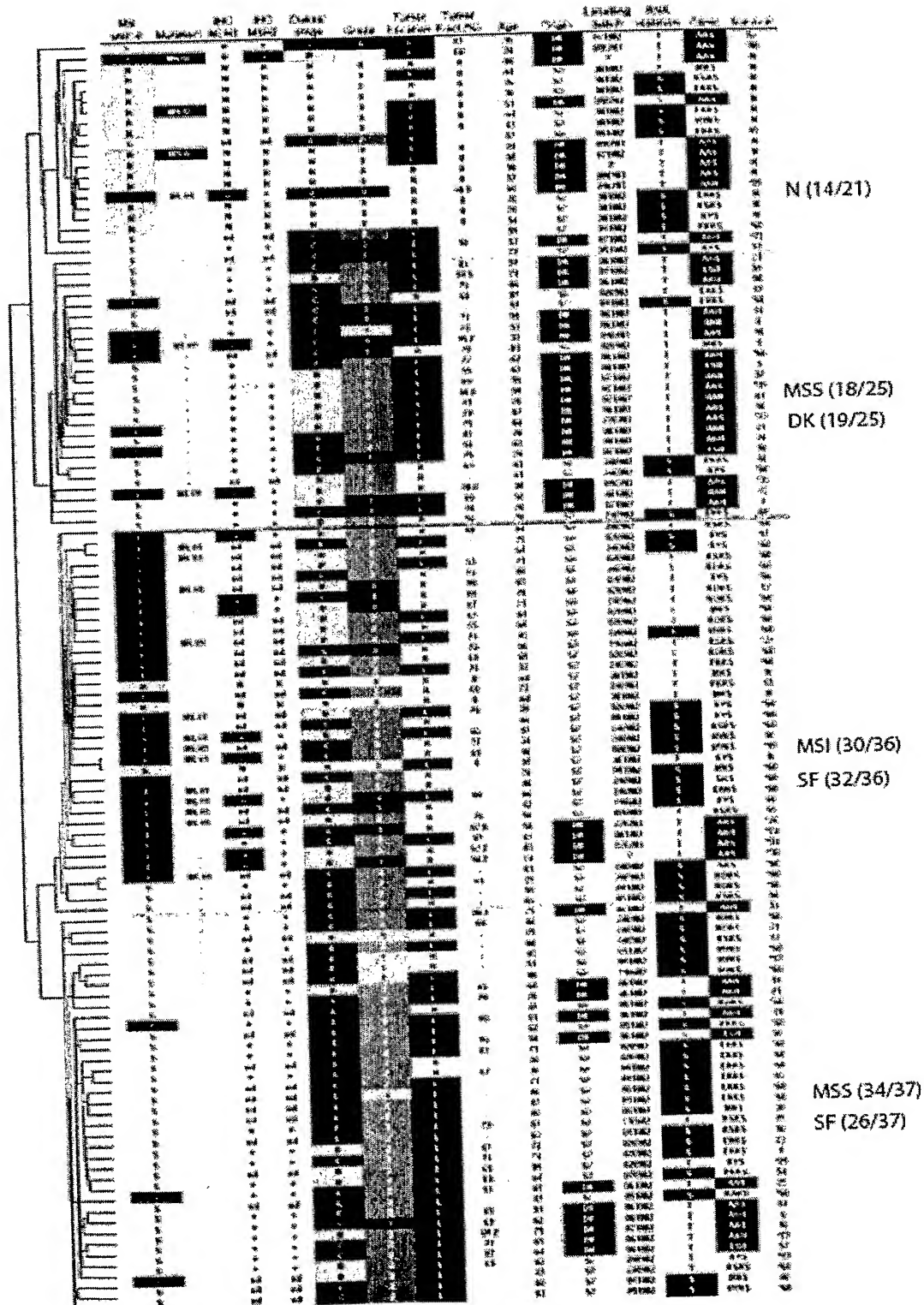
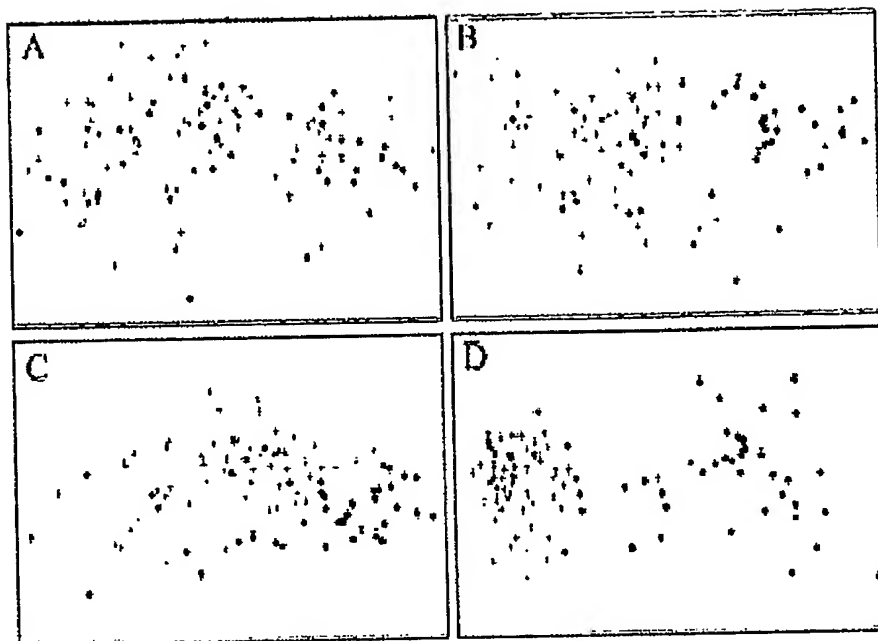


Fig. 7/10



26 NOV. 2004

PVS

Fig. 8/10

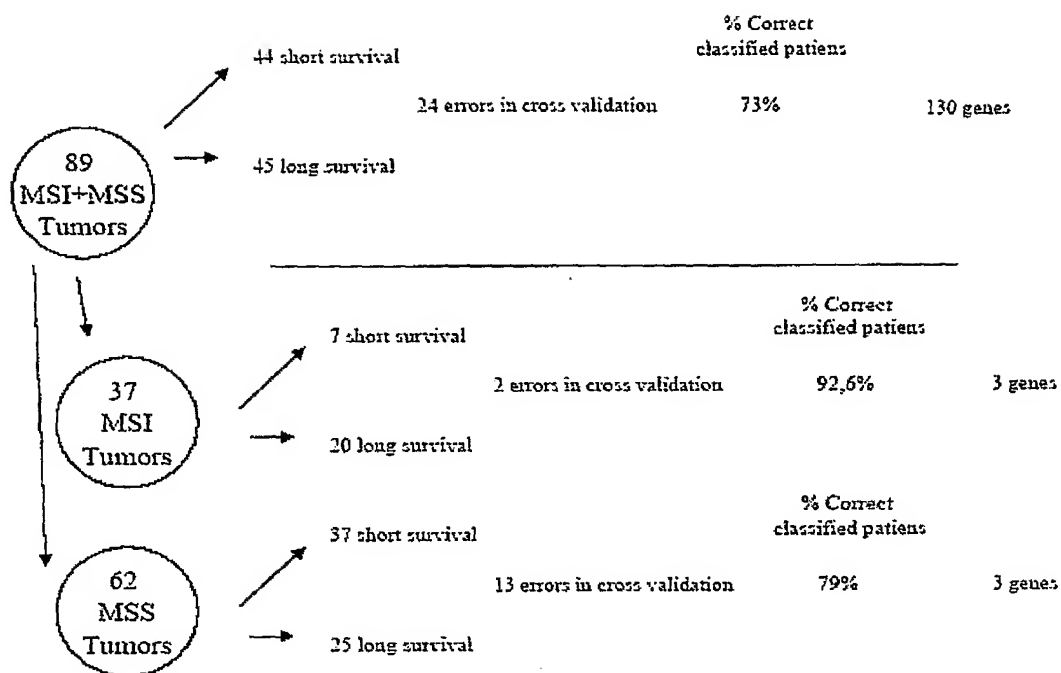
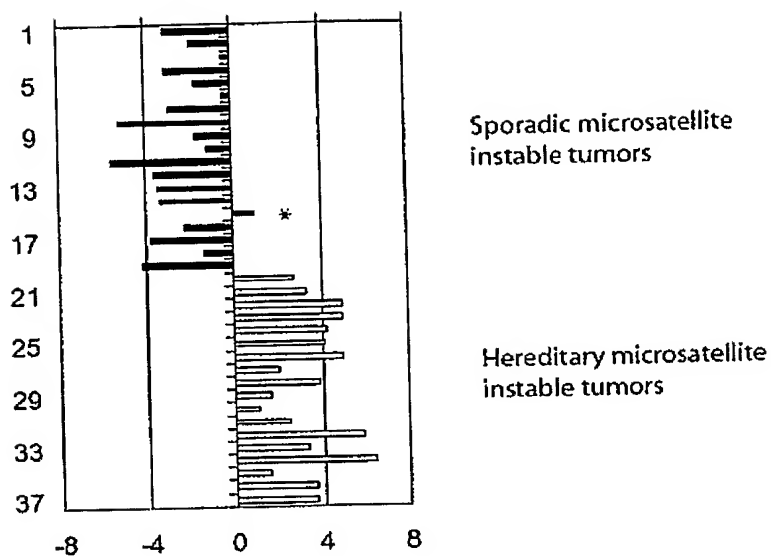


Fig. 9/10

A



B

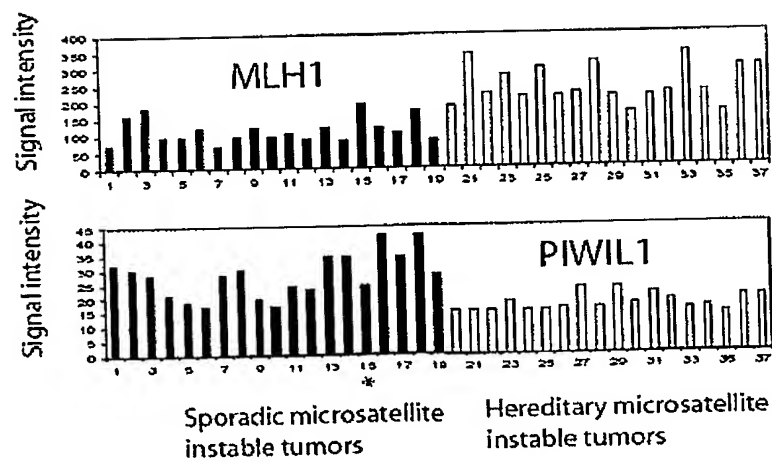
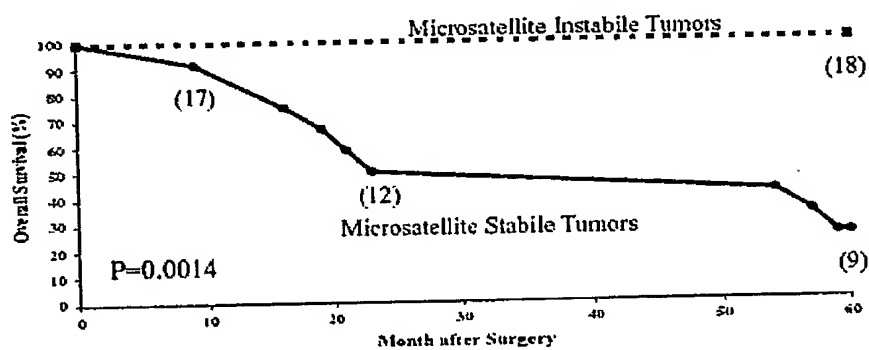
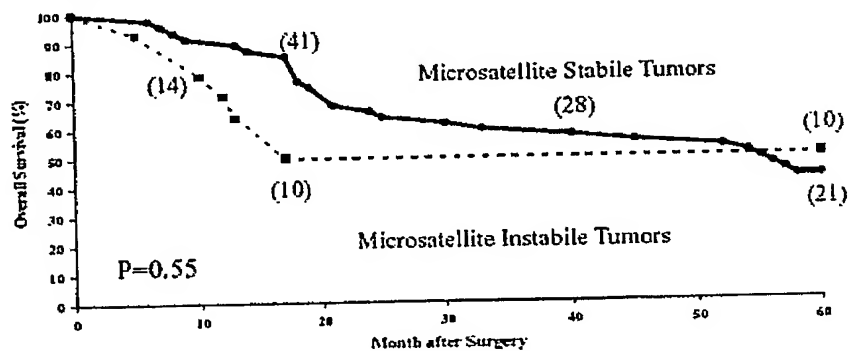


Fig. 10/10

## A Patients with Dukes' B Colon Cancer (No adjuvant Chemotherapy)



## B Patients with Dukes' C Colon Cancer (Adjuvant Chemotherapy)



SEQUENCE LISTING

Modtaget

26 NOV. 2004

PVS

<110> Aros Applied Biotechnology ApS

<120> Classification of Colon Cancer

<130> P949DK00

<160> 139

<170> PatentIn version 3.1

<210> 1

<211> 1237

<212> DNA

<213> NM\_002985.2| Homo sapiens chemokine (C-C motif) ligand 5 (CCL5), mRNA

```

<400> 1
gctgcagagg attcctgcag aggatcaaga cagcacgtgg acctcgacaca gcctctccca 60
caggtaccat gaaggtctcc gcggcagccc tcgctgtcat cctcattgct actgccctct 120
gcgctcctgc atctgcctcc ccatattcct cggacaccac accctgctgc ttgcctaca 180
ttgccccgcc actgccccgt gccacatca aggagtattt ctacaccagt ggcaagtgt 240
ccaaccagc agtcgtcttt gtcacccgaa agaaccgcca agtgtgtgcc aaccagaga 300
agaaatgggt tcgggagtac atcaactctt tggagatgag ctaggatgga gagtcttga 360
acctgaactt acacaaattt gcctgtttct gcttgctctt gtcctagctt gggaggcttc 420
ccctcactat cctacccac ccgctccttg aagggccag attctaccac acagcagcag 480
ttacaaaaac cttccccagg ctggacgtgg tggctcacgc ctgtaatccc agcactttgg 540
gaggccaagg tgggtggatc acttgaggtc aggagttcga gaccagcctg gccaacatga 600
tgaaaccca tcttactaa aaatacaaaa aattagccgg gcgtggtagc gggcgctgt 660
agtcccagct actcgggagg ctgaggcagg agaatggcgt gaaccggga ggcggagctt 720
gcagtgagcc gagatcgcg cactgcactc cagcctgggc gacagagcga gactccgtct 780
caaaaaaaaa aaaaaaaaaa aaaatacaaaa aattagccgg gcgtggtagc ccacgcctgt 840
aatcccagct actcgggagg ctaaggcagg aaaattgttt gaaccagga ggtggaggct 900
gcagtgagct gagattgtgc cacttcactc cagcctgggt gacaaagtga gactccgtca 960

```

caacaacaac aacaaaaagc ttccccaact aaagcctaga agagcttctg aggcgctgct 1020  
 ttgtcaaaag gaagtctcta ggttctgagc tctggctttg ccttggcttt gccagggctc 1080  
 tgtgaccagg aaggaagtca gcatgcctct agaggcaagg aggggaggaa cactgcactc 1140  
 ttaagcttcc gccgtctcaa cccctcacag gagcttactg gcaaaccatga aaaatcggct 1200  
 taccattaaa gtcttcaatg caaccataaa aaaaaaa 1237

<210> 2

<211> 2884

<212> DNA

<213> NM\_004184.3| Homo sapiens tryptophanyl-tRNA synthetase (WARS),  
 transcript variant 1, mRNA

<400> 2  
 tcgattctca agagggtttc attggtctca acctggcccc ccaggcaacc caccctgat 60  
 tggacagtct catcaagaag gttggtcaag agctcaagtg tttctgagaa tctgggtgat 120  
 ttataagaaa cccttagctg aatgcagggg ggggagaacg aaagacaaaa gcatcttttt 180  
 tcagaaggga aactgaaaga aagaggggaa gagtattaaa gaccatttct ggctgggcag 240  
 ggcactctca gcagctcaac tgcccagcgt gaccagtggc cacctctgca gtgtcttcca 300  
 caacctggtc ttgactcgtc tgctgaacaa atcctctgac ctgagccggg ctgtgaacgt 360  
 agttcctgag agatagcaaa catgccccac agtgagcccg catctctgct ggagctgttc 420  
 aacagcatcg ccacacaagg ggagctcgta aggtccctca aagcgggaaa tgcgtcaaag 480  
 gatgaaattg attctgcagt aaagatgttg gtgtcattaa aaatgagcta caaagctgcc 540  
 gcgggggagg attacaaggc tgactgtcct ccagggaacc cagcacctac cagtaatcat 600  
 ggcccagatg ccacagaagc tgaagaggat tttgtggacc catggacagt acagacaagc 660  
 agtgcaaaaag gcatagacta cgataagctc attgttcggt ttggaagtag taaaattgac 720  
 aaagagctaa taaaccgaat agagagagcc accggccaaa gaccacacca cttcctgcgc 780  
 agaggcatct tcttctcaca cagagatatg aatcaggttc ttgatgccta tgaaaataag 840  
 aagccatttt atctgtacac gggccggggc cctcttctg aagcaatgca tgtaggtcac 900  
 ctcatccat ttattttcac aaagtggctc caggatgtat ttaacgtgcc cttgggtcatc 960  
 cagatgacgg atgacgagaa gtatctgttg aaggacctga ccctggacca ggcctatagc 1020  
 tatgctgtgg agaattgcaa ggacatcatc gcctgtggct ttgacatcaa caagactttc 1080  
 atattctctg acctggacta catgggatg agctcagggt tctacaaaaa tgtggtgaag 1140  
 attcaaaagc atgttacctt caaccaagtg aaaggcattt tcggcttcac tgacagcgac 1200  
 tgcattggga agatcagttt tcctgccatc caggctgctc cctccttcag caactcattc 1260  
 ccacagatct tccgagacag gacggatatc cagtgcctta tcccatgtgc cattgaccag 1320



gatccttact ttagaatgac aagggacgtc gccccagga tcggctatcc taaaccagcc	1380
ctgtgcact ccaccttctt cccagccctg cagggcgccc agaccaaaat gagtgccagc	1440
gaccccaact cctccatctt cctcaccgac acggccaagc agatcaaaac caaggtcaat	1500
aagcatgcgt tttctggagg gagagacacc atcgaggagc acaggcagtt tgggggcaac	1560
tgtgatgtgg acgtgtcttt catgtacctg accttcttcc tcgaggacga cgacaagctc	1620
gagcagatca ggaaggatta caccagcgga gccatgtcga ccggtgagct caagaaggca	1680
ctcatagagg ttctgcagcc cttgatcgca gagcaccagg cccggcgcaa ggaggtcacg	1740
gatgagatag tgaaagagtt catgactccc cggaagctgt ccttcgactt tcagtagcac	1800
tcgttttaca tatgcttata aaagaagtga tgtatcagta atgtatcaat aatcccagcc	1860
cagtcaaagc accgccacct gtaggcttct gtctcatggt aattactggg cctggcctct	1920
gtaagcctgt gtatgtttatc aatactgttt cttcctgtga gttccattat ttctatctct	1980
tatgggcaaa gcattgtggg taattggtgc tggctaacat tgcattggtcg gatagagaag	2040
tccagctgtg agtctctccc caaagcagcc ccacagtgga gcctttggct ggaagtccat	2100
gggccaccct gttcttgtcc atggaggact ccgagggttc caagtatact cttaagaccc	2160
actctgttta aaaatatata ttctatgtat gcgtatatgg aattgaaatg tcattattgt	2220
aacctagaaa gtgctttgaa atattgatgt ggggaggttt attgagcaca agatgtattt	2280
cagcccatgc cccctcccaa aaagaaattg ataagtaaaa gcttcgttat acatttgact	2340
aagaaatcac ccagctttaa agctgctttt aacaatgaag attgaacaga gttcagcaat	2400
tttgattaaa ttaagacttg ggggtgaaac tttccagttt actgaactcc agaccatgca	2460
tgtagtccac tccagaaatc atgctcgctt cccttggcac accagtgttc tcctgccaaa	2520
tgaccctaga ccctctgtcc tgcagagtca ggggtgcttt tcccctgact gtgtccgatg	2580
ccaaggagtc ctggcctccg cagatgcttc attttgacct ttggctgcag tggaagtcag	2640
cacagagcag tgccctggct gtgtccctgg acgggtggac ttagctaggg agaaagtcga	2700
ggcagcagcc ctcgaggccc tcacagatgt ctaggcaggc ctcatttcat cacgcagcat	2760
gtgcaggcct ggaagagcaa agccaaatct caggaagtc cttggttgat gtatctgggt	2820
ctcctctgga gcaactctgcc ctcctgtcac ccagtagagt aaataaactt ccttggctcc	2880
tgct	2884

<210> 3

<211> 1012

<212> DNA

<213> NM\_006263.2| Homo sapiens proteasome (prosome, macropain) activator subunit 1 (PA28 alpha) (PSME1), transcript variant 1, mRNA

<400> 3

aggcggagct ggggtgcgagc gccctaccgc tttcgctttc ctttcgcggt gcccactcca	60
ctccttgtgc ggcgctaggc cccccgtccc ggtcatggcc atgctcaggg tccagcccga	120
ggcccaagcc aaggtggatg tgtttcgtga agacctctgt accaagacag agaacctgct	180
cgggagctat ttccccaaga agatttctga gctggatgca tttttaaagg agccagctct	240
caatgaagcc aacttgagca atctgaaggc cccattggac atcccagtgc ctgatccagt	300
caaggagaaa gagaaagagg agcgggaagaa acagcaggag aaggaagaca aggatgaaaa	360
gaagaagggg gaggatgaag acaaaggctc tccctgtggc ccagtgaact gcaatgaaaa	420
gatcgtggtc cttctgcagc gcttgaagcc tgagatcaag gatgtcattg agcagctcaa	480
cctggtcacc acctggttgc agctgcagat acctcggatt gaggatggta acaattttgg	540
agtggctgtc caggagaagg tgtttgagct gatgaccagc ctccacacca agctagaagg	600
cttccacact caaatctcta agtatttctc tgagcgtggg gatgcagtga ctaaagcagc	660
caagcagccc catgtgggtg attatcgga gctgggtgcac gagctggatg aggcagagta	720
ccgggacatc cggctgatgg tcatggagat ccgcaatgct tatgctgtgt tatatgacat	780
catcctgaag aacttcgaga agctcaagaa gcccagggga gaaacaaagg gaatgatcta	840
ttgagagccc tctctcccat tctgtgatga gtacagcaga gaccttcctg ctttttactg	900
gggactccag attttcccca aacttgcttc tgttgagatt tttccctcac cttgcctctc	960
aggcacaata aatatagtta taccactgcc catcaaaaaa aaaaaaaaaa aa	1012

<210> 4

<211> 983

<212> DNA

<213> NM\_004335.2| Homo sapiens bone marrow stromal cell antigen 2 (BST2), mRNA

<400> 4	
gtggaattca tggcatctac ttcgtatgac tattgcagag tgcccatgga agacggggat	60
aagcgtgta agcttctgct ggggatagga attctggtgc tcctgatcat cgtgattctg	120
ggggtgccct tgattatctt caccatcaag gccaacagcg aggcctgccg ggacggcctt	180
cgggcagtga tggagtgtcg caatgtcacc catctcctgc aacaagagct gaccgaggcc	240
cagaagggct ttcaggatgt ggaggcccag gccgccacct gcaaccacac tgtgatggcc	300
ctaattggctt ccctggatgc agagaaggcc caaggacaaa agaaagtgga ggagcttgag	360
ggagagatca ctacattaaa ccataagctt caggacgcgt ctgcagaggt ggagcgactg	420
agaagagaaa accaggtctt aagcgtgaga atcgcggaca agaagtacta cccagctcc	480
caggactcca gctccgctgc ggcgccccag ctgctgattg tgctgctggg cctcagcgct	540
ctgctgcagt gagatcccag gaagctggca catcttgga ggtccgctct gctcggcttt	600

tcgcttgaac attcccttga tctcatcagt tctgagcggg tcatggggca acacgggtag	660
cgggggagagc acggggtagc cggagaaggg cctctggagc aggtctggag gggccatggg	720
gcagtcctgg gtgtggggac acagtcgggt tgacccaggg ctgtctccct ccagagcctc	780
cctccggaca atgagtcccc cctcttgtct cccaccctga gattgggcat ggggtgcggt	840
gtggggggca tgtgctgcct gttgttatgg gttttttttg cggggggggg tgcttttttc	900
tgggggtcttt gagctccaaa aaataaacac ttcctttgag ggagagcaaa aaaaaaaaaa	960
aaaaaaaaa aaaaaaaaaa aaa	983

<210> 5

<211> 1260

<212> DNA

<213> NM\_004223.3| Homo sapiens ubiquitin-conjugating enzyme E2L 6 (UBE2L6), transcript variant 1, mRNA

<400> 5	
gggggtgggg tccccggggc ggggcggggc gcgctgtgtc gcgggtcgga gctcggtcct	60
gctggaggcc acgggtgcc acaactcgggt cccgacatga tggcgagcat gcgagtgggtg	120
aaggagctgg aggatcttca gaagaagcct ccccatatac tgcggaacct gtccagcgat	180
gatgccaatg tcctggtgtg gcacgctctc ctctaccgg accaacctcc ctaccacctg	240
aaagccttca acctgcgcac cagcttcccc cgggagtatc cgttcaagcc tcccatgatc	300
aaattcacia ccaagatcta ccaccccaac gtggacgaga acggacagat ttgcctgccc	360
atcatcagca gtgagaactg gaagccttgc accaagactt gccaaagtcct ggaggccctc	420
aatgtgctgg tgaatagacc gaatatcagg gagccccctgc ggatggacct cgctgacctg	480
ctgacacaga atccggagct gttcagaaag aatgccgaag agttcacctc ccgattcgga	540
gtggaccggc cctcctaact catgttctga ccctctgtgc actggatcct cggcatagcg	600
gacggacaca cctcatggac tgaggccaga gccccctgtg gccattccc cattcatttt	660
tcccttctta ggttgtagt cattagtttg tgtgtgtgtg tgggtggagg aaggagagcta	720
tgagtgtgtg tgttggtat ggactcactc ccagggttcac ctggccacag gtgcaccctt	780
cccacaccct ttacattccc cagagccaag ggagtttaag tttgcagtta caggccagtt	840
ctccagctct ccattctaga gagacaggtc accttgacagg cctgcttgca ggaaatgaat	900
ccagcagcca actcgaatcc ccctagggct caggcactga gggcctgggg acagtggagc	960
atatgggtgg gagacagatg gagggtagcc tatttacaac tgagtcagcc aagccactga	1020
tgggaatata cagatttagg tgctaaaccg tttattttcc acggatgagt cacaatctga	1080
agaatcaaac ttccatcctg aaaatctata tgtttcaaaa ccacttgcca tcctgtaga	1140
ttgccagttc ctgggaccag gcctcagact gtgaagtata taccctccag cattcagtc	1200

agggggagcc acggaaccca tgttcttgct taagccatta aagtcagaga tgaattctgg 1260

<210> 6

<211> 3799

<212> DNA

<213> NM\_003488.2| Homo sapiens A kinase (PRKA) anchor protein 1 (AKAP1), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA

<400> 6  
ctgtgttcca cccgcctggg ctagcacgtg ggggagctgc ggaagcgcgg cgctgcgggc 60  
cgggccgcgg ggcacagccg ggggccggcg gcggcgcgcg gactccgcac cccgcacccc 120  
gatggtagcc gaggagctgg tgtaattact tcaagcctcc aggatggcaa tccagttccg 180  
ttcgtcttcc cccttgccat tgcctgggat gctggcgctc ctcggctggt ggtggttttt 240  
ctctcgtaaa aaaggccatg tcagcagcca tgatgagcag cagggtggagg ctggtgctgt 300  
gcagctgagg gctgaccctg ccatcaagga acctctcccc gtggaagacg tctgtcccaa 360  
agtagtgctc acacccccca gtgtcacaga gcctccagaa aaggaactgt ccaccgtgag 420  
caagctgcct gcagagcccc cagcattgct ccagacacac ccaccttgcc gaagatcaga 480  
gtcctcgggc attcttccca acaccacaga catgagattg cgaccaggaa cacgcagaga 540  
tgacagtaca aagctggagc tagccctgac aggtggtgaa gccaaatcga ttcctctaga 600  
gtgccccctt tcatcccca aaggtgtact attctccagc aaatcagctg aggtgtgtaa 660  
gcaagattcc cccttcagca gggtgccaag gaaggtccag ccaggctacc ccgtagtccc 720  
cgcagagaag cgtagctctg gggagagggc aagagagaca ggtggggccg aagggactgg 780  
tgatgccgtg ttgggggaaa aggtgcttga agaagctctg ttgtctcggg agcatgtctt 840  
ggaattggag aacagcaagg gccccagcct ggcctcttta gagggggaag aagataaggg 900  
gaagagcagc tcatcccagg tgggtggggc agtgcaggag gaagagtatg tagcagagaa 960  
gttgccaagt aggttcacat agtcggctca cacagagctg gcaaaggacg atgcggcgcc 1020  
agcaccacca gtcgcagacg ccaaagccca ggatagaggt gtcgagggag aactgggcaa 1080  
tgaggagagc ttggatagaa atgaggaggg cttggataga aatgaggagg gcttggatag 1140  
aaatgaggag agcttggata gaaatgagga gggcttggat agaaatgagg agattaagcg 1200  
ggctgccttc cagataatct cccaagtgat ctcaagaagca accgaacagg tgctggccac 1260  
cacggttggc aaggttgtag gtcgtgtgtg tcaggccagt cagctccaag ggcagaagga 1320  
agagagctgt gtcccagttc accagaaaac tgtcttgggc ccagacactg cggagcctgc 1380  
cacagcagag gcagctgttg ccccgccgga tgctggcctc cccttgccag gcctaccagc 1440  
agagggctca ccaccacca agacctacgt gagctgcctg aagagccttc tgtccagccc 1500  
caccaaggac agtaagccaa atatctctgc acaccacatc tccctggcct cctgcctggc 1560

actgaccacc cccagtgaag agttgccgga ccgggcaggc atcctggtgg aagatgccac	1620
ctgtgtcacc tgcattgtcag acagcagcca aagtgtccct ttggtggcct ctccaggaca	1680
ctgtctcagat tctttcagca cttcagggtc tgaagactct tgcacagaga ccagctcgag	1740
ccccagggac aaggccatca ccccgccact gccagaaagt actgtgccct tcagcaatgg	1800
gggtgctgaag ggggagttgt cagacttggg ggctgaggat ggatggacca tggatgcgga	1860
agcagatcat tcaggagggt ctgacaggaa cagcatggat tccgtggata gctgttgag	1920
tctcaagaag actgagagct tccaaaatgc ccaggcaggc tccaacccta agaaggtcga	1980
cctcatcatc tgggagatcg aggtgccaaa gcacttagtc ggtcggctaa ttggcaagca	2040
ggggcgctat gtgagttttc tgaagcaaac atctggtgcc aagatctaca tttcaaccct	2100
gccttacacc cagagcgctc agatctgcca catagaaggc tctcaacatc atgtagacaa	2160
agcgtgaac ttgattggga agaagttcaa agagctgaac ctaccaata tctacgctcc	2220
cccattgcct tcaactggcag tgccttctct gccgatgaca tccctggctca tgctgcctga	2280
tggcatcacc gtggagggtca ttgtggtcaa ccagggtcaat gccgggcacc tgttcgtgca	2340
gcagcacaca caccctacct tccacgcgct gcgcagcctc gaccagcaga tgtacctctg	2400
ttactctcag cctggaatcc ccaccttgcc caccctcagt gaaataacgg tcatctgtgc	2460
cgccccctggt gcggacgggg cctggtggcg agcccaagtg gttgcctcct acgaggagac	2520
caacgaagtg gagattcgat acgtggacta cggcggatat aagaggggtga aagtagacgt	2580
gctccggcaa atcaggtctg actttgtcac cctgccgttt caggagcag aagtccttct	2640
ggacagtgtg atgccccctgt cagacgatga ccagttttca ccggaagcag atgccgccat	2700
gagcgagatg acggggaata cagcactgct tgctcagggtg acaagttaca gtccaactgg	2760
tcttcctctg attcagctgt ggagtgtggt tggagatgaa gtggtgttga taaaccggtc	2820
cctggtggag cgaggccttg cccagtgggt agacagctac tacacaagcc tttgaccccc	2880
atgctgcttc ctgagagtct ttttttgac tggtgaaatt gggcttggca ctcaagtcaa	2940
agatgaacat cggaataaca aacattgtcc tctccagaaa gtcctttctt tatccatact	3000
gtagtcctat tgagaagaca tttcgtctct gagaaaaaag gatggaacta tgggttctct	3060
tcgcaaagcc aaaggatagt gtttaacaag ccagctggct taccctggtt ctcaactgtt	3120
taaaaaaaaa aaaaaaagg aatagaaaca gtttcaacca gattgtccta tccccctgt	3180
tccattcccc tcttcttctt tctatctcct tccccggcaa aaaccaaaca aactggcaga	3240
caggccaggg atgtatgttg cttgcttgag agggtttctt ttacttcaaa atctttcttc	3300
agggagcaag acatgaactg actaattggt atccactact tgtacagctt acataaatga	3360
gttgatgata tttaaccagt ttttataaac ttcatttagg tctctaaaca cagacttttt	3420
aaattgcaac tgtaaataatg aaatgggtcat cacatctgac cttgggtcagt ggggagggga	3480
actggtatcc tgccaagcct gggtgtaatt tgtaaccatt ttctatttgt gcaaactctg	3540
taaatatgtg tttaaacaaa tgtaatatat tgtacaagat aactgggaga acaaagggaa	3600

ctcaagattc ttccagccac atgtcacctg taggtagaag taaactctgc agtgcagctt 3660  
ctgctcttgg cccctctggc cagggcccct gtggcttcct gcacactgga caggtgactg 3720  
tatggtagag actgtgatct ggggaactttt tgctgtacaa atctgtttta aaaaaaaaaa 3780  
aagtaactca ttgaattaa 3799

<210> 7

<211> 829

<212> DNA

<213> NM\_002818.2| Homo sapiens proteasome (prosome, macropain) activator subunit 2 (PA28 beta) (PSME2), mRNA

<400> 7  
tggggagtga aagcgaaagc ccgggcgact agccgggaga ccagagatct agcgactgaa 60  
gcagcatggc caagccgtgt ggggtgccc tgagcgggga agcccgcaa caggtggagg 120  
tcttcagaca gaatcttttc caggaggctg aggaattcct ctacagattc ttgccacaga 180  
aaatcatata cctgaatcag ctcttgcaag aggactccct caatgtggct gacttgactt 240  
ccctccgggc cccactggac atccccatcc cagaccctcc acccaaggat gatgagatgg 300  
aaacagataa gcaggagaag aaagaagtcc ataagtgtgg atttctccct gggaatgaga 360  
aagtcctgtc cctgcttgcc ctggttaagc cagaagtctg gactctcaa gagaaatgca 420  
ttctggtgat tacatggatc caacacctga tccccaagat tgaagatgga aatgattttg 480  
gggtagcaat ccaggagaag gtgctggaga gggatgaatgc cgtcaagacc aaagtggaag 540  
ctttccagac aaccatttcc aagtacttct cagaacgtgg ggatgctgtg gccaaggcct 600  
ccaaggagac tcatgtaatg gattaccggg ccttggtgca tgagcgagat gaggcagcct 660  
atggggagct cagggccatg gtgctggacc tgagggcctt ctatgctgag ctttatcata 720  
tcatcagcag caacctggag aaaattgtca acccaaaggg tgaagaaaag ccatctatgt 780  
actgaaccg ggactagaag gaaaataaat gatctatatg ttgtgtgga 829

<210> 8

<211> 2974

<212> DNA

<213> NM\_004363.1 Homo sapiens carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5), mRNA

<400> 8  
ctcagggcag agggaggaag gacagcagac cagacagtca cagcagcctt gacaaaacgt 60  
tcctggaact caagctcttc tccacagagg aggacagagc agacagcaga gaccatggag 120

tctccctcgg cccctcccca cagatggtgc atcccctggc agaggctcct gctcacagcc	180
tcactttctaa ctttctggaa cccgcccacc actgccaaagc tcactattga atccacgccg	240
ttcaatgtcg cagaggggaa ggaggtgctt ctacttgtcc acaatctgcc ccagcatctt	300
tttggttaca gctggtacaa aggtgaaaga gtggatggca accgtcaaata tataggatat	360
gtaataggaa ctcaacaagc taccacaggg cccgcataca gtggctcgaga gataatatac	420
cccaatgcat ccctgctgat ccagaacatc atccagaatg acacaggatt ctacacccta	480
cacgtcataa agtcagatct tgtgaatgaa gaagcaactg gccagttccg ggtatacccg	540
gagctgcccc agccctccat ctccagcaac aactccaaac ccgtggagga caaggatgct	600
gtggccttca cctgtgaacc tgagactcag gacgcaacct acctgtggtg ggtaaacaat	660
cagagcctcc cggctcagtc caggctgcag ctgtccaatg gcaacaggac cctcactcta	720
ttcaatgtca caagaaatga cacagcaagc tacaaatgtg aaaccagaa cccagtgaat	780
gccaggcgca gtgattcagt catcctgaat gtcctctatg gcccggatgc cccaccatt	840
tcccctctaa acacatctta cagatcaggg gaaaatctga acctctcctg ccacgcagcc	900
tctaaccac ctgcacagta ctcttggttt gtcaatggga ctttccagca atccacccaa	960
gagctcttta tccccaacat cactgtgaat aatagtggat cctatacgtg ccaagcccat	1020
aactcagaca ctggcctcaa taggaccaca gtcacgacga tcacagtcta tgcagagcca	1080
cccaaaccct tcatcaccag caacaactcc aaccccgtagg aggatgagga tgctgtagcc	1140
ttaacctgtg aacctgagat tcagaacaca acctacctgt ggtgggtaaa taatcagagc	1200
ctcccggta gtcccaggct gcagctgtcc aatgacaaca ggaccctcac tctactcagt	1260
gtcacaagga atgatgtagg acctatgag tgtggaatcc agaacgaatt aagtgttgac	1320
cacagcgacc cagtcacct gaatgtctc tatggcccag acgacccac catttcccc	1380
tcatacacct attaccgtcc aggggtgaac ctacgctct cctgccatgc agcctctaac	1440
ccacctgcac agtattcttg gctgattgat gggaacatcc agcaacacac acaagagctc	1500
tttatctcca acatcactga gaagaacagc ggactctata cctgccaggc caataactca	1560
gccagtggcc acagcaggac tacagtcaag acaatcacag tctctgcgga gctgcccag	1620
ccctccatct ccagcaacaa ctccaaacc gtggaggaca aggatgctgt ggccttcacc	1680
tgtgaacctg aggtcagaa cacaacctac ctgtgggtggg taaatggtca gagcctcca	1740
gtcagtcca ggctgcagct gtccaatggc aacaggacct tcactctatt caatgtcaca	1800
agaaatgacg caagagccta tgtatgtgga atccagaact cagtgaagtgc aaaccgcagt	1860
gacccagtca ccctggatgt cctctatggg ccggacaccc ccatcatttc cccccagac	1920
tcgtcttacc tttcgggagc gaacctcaac ctctcctgcc actcggcctc taacctatcc	1980
ccgcagtatt cttggcgat caatgggata ccgcagcaac acacacaagt tctctttatc	2040
gccaaaatca cgccaaataa taacgggacc tatgcctgtt ttgtctctaa cttggctact	2100
ggccgcaata attccatagt caagagcatc acagtctctg catctggaac ttctcctggt	2160

ctctcagctg gggccactgt cggcatcatg attggagtgc tggttggggt tgctctgata	2220
tagcagccct ggtgtagttt cttcatttca ggaagactga cagttgtttt gcttcttcct	2280
taaagcattt gcaacagcta cagtctaaaa ttgcttcttt accaaggata ttacagaaa	2340
agactctgac cagagatcga gaccatccta gccaacatcg tgaaaccca tctctactaa	2400
aaatacaaaa atgagctggg cttggtggcg cgcacctgta gtcccagtta ctctgggaggc	2460
tgaggcagga gaatcgcttg aaccggggag gtggagattg cagtgaagccc agatcgacc	2520
actgcactcc agtctggcaa cagagcaaga ctccatctca aaaagaaaag aaaagaagac	2580
tctgacctgt actcttgaat acaagtttct gataccactg cactgtctga gaatttccaa	2640
aactttaatg aactaactga cagcttcctg aaactgtcca ccaagatcaa gcagagaaaa	2700
taattaattt catgggacta aatgaactaa tgaggattgc tgattcttta aatgtcttgt	2760
ttcccagatt tcaggaaact ttttttcttt taagctatcc actcttacag caatttgata	2820
aaatatactt ttgtgaacaa aaattgagac atttacattt tctccctatg tggcgcctcc	2880
agacttgga aactattcat gaatatttat attgtatggt aatatagtta ttgcacaagt	2940
tcaataaaaa tctgctcttt gtataacaga aaaa	2974

<210> 9

<211> 5028

<212> DNA

<213> NM\_005766.2| Homo sapiens FERM, RhoGEF (ARHGEF) and plectstrin domain protein 1 (chondrocyte-derived) (FARP1), transcript variant 1, mRNA

<400> 9	
gcggccgctg cccgctttgc gccgctcctc cctgcgcgag tagcgcctggc cccggcgctcg	60
aggcgggccat ggcgacccgg agcccgctcc ccacccaccc cgctctctcc gccctcccct	120
ccgccccgcg ccacctttga tggctcggac ctacgcccgc caccgccagc cctgctcgcg	180
cgccccgccc gccgcccggc gcgggtatta atagccggcg ccgcccgcgc ctcggccgccc	240
gggggcttgg gagccgccga tcccggagcc cgagccggga gagggagccg ccgcagccgc	300
cggcgctgtg gagatattct ctaagccgct ttcacatgag gagaaataga gcagaggccg	360
acccagagat cagcactggg ggccccggaa aattcgggga tcagtacctt ggaacgtgga	420
cagaagccgc cccaacacc ttcaggaaaa ctctgtgtcca tcaaaatcca gatgctggat	480
gacaccagag aggcatttga agttccacaa agagctcctg ggaagggtgct gctggatgca	540
gtttgcaacc acctcaacct cgtggaaggt gactattttg gcctcgagtt tcctgatcac	600
aaaaagatca cgggtgtggc ggatctccta aaaccattg tgaaacagat tagaaggcca	660
aagcacgttg ttgttaagtt tgtggtgaaa ttctttccgc ctgaccacac acaactccaa	720
gaagaactca caaggtagct gtccgcgctg cagggtgaagc aggacttggc tcaaggcagg	780



ttgacgtgta atgacaccag cgcagctctc ttgatttcac acattgtgca atctgagatt	840
ggggattttg atgaagcctt ggacagagag cacttagcaa aaaataaata catacctcag	900
caagacgcac tagaggacaa aatcgtggaa tttcaccata accacattgg acaaacacca	960
gcagaatcag atttccagct cctagagatt gcccgtcggc tagagatgta tggaatccgg	1020
ttgcacccgg ccaaggacag ggaaggcacg aagatcaatc tggccgttgc caacacggga	1080
attctagtgt ttcagggttt cactaagatc aatgccttca actgggcca a ggtgcggaag	1140
ctgagcttca agaggaagcg ctttctcatc aagctccggc cagatgcca tagtgcgta	1200
caggatacct tggaattcct gatggccagt cgggatttct gcaagtcctt ctggaaaatc	1260
tgtgttgaa atcatgcctt ctttagactt tttgaagagc ccaaaccaaa gcccaagccc	1320
gtcctcttta gccgggggtc atcatttcgg ttcagtggc ggactcagaa gcaggttctc	1380
gactatgtta aagaaggagg acataagaag gtgcagttt aaaggaagca cagcaagatt	1440
cattctatcc ggagccttgc ttcacagcct acagaactga attcggaagt gctggagcag	1500
tctcagcaga gcaccagcct tacatttgga gaaggtgccg aatctccagg gggccagagc	1560
tgccggcgag gaaaggaacc gaaggtttcc gccggggagc cggggtcgca cccgagccct	1620
gcgcccagga gaagccccgc gggtaacaag caggcggacg gagccgcctc ggcgcccacg	1680
gaggaagagg aggaggtcgt taaggatagg acccagcaga gtaaacctca gccccgcag	1740
ccaagcacag gctccctgac tggcagtcct cacctttccg agctgtctgt gaactcgag	1800
gggggagtg cccctgccaa cgtgaccttg tctcccaacc tgagccccga caccaagcag	1860
gcctctccct tgatcagccc gctgctgaat gaccaggcct gccccggac ggacgatgag	1920
gatgagggcc ggaggaagag attcccaact gataaagcgt acttcatagc taaggaagt	1980
tctaccaccg agcgaacata tctgaaggat ctcgaagtta tcacttcgtg gtttcagagc	2040
acagtgcag aagaggacgc catgccgga gactgaaaa gtctcatatt cccgaatttt	2100
gaacctttgc acaaatttca tactaatttt ctcaaggaaa ttgagcaacg acttgccctg	2160
tgggaaggcc gctcaaatgc ccaaatcaga gattaccaa gaatcggcga tgtcatgctg	2220
aagaacattc agggcatgaa gcacctggcg gctcacctgt ggaagcacag cgaggccttg	2280
gaggccctgg agaatggaat caagagctcc cggcggtgg agaacttctg cagagacttt	2340
gagctgcaga aggtgtgtta cctaccgctc aacaccttcc tctgcgggc actgcaccgg	2400
ctcatgcact acaagcaggt cctggagcgg ctgtgcaaac accacccgcc gagccacgcc	2460
gacttcaggg actgcccagc cgctttggca gagatcacgg agatggtggc acagctccac	2520
ggtacgatga tcaagatgga gaatttccag aagctgcacg aactcaagaa agatttgatt	2580
ggcattgaca atcttggtt tccgggaagg gagttcatcc gtctgggcag cctcagcaag	2640
ctctcgggga aggggctcca gcagcgcatg ttcttctgt tcaacgacgt cctgctatac	2700
acgagccggg ggctgacggc ctccaatcag tttaaagtcc acgggcagct cccgctctat	2760

ggcatgacga ttgaggagag cgaagacgag tggggggtgc cccactgcct gaccctccgg	2820
ggccagcggc agtccatcat cgtggccgcc agttctcggt ccgagatgga gaagtgggtt	2880
gaggacatcc agatggccat tgacctggcg gagaagagca gcagccccgc ccctgagttc	2940
ctggccagca gccccctga caacaagtcc cctgatgaag ccaccgcggc tgaccaggag	3000
tcagaggatg acctgagcgc ctcgcgcaca tcgctggagc gccaggcccc gcaccgcggc	3060
aacacaatgg tgcacgtgtg ctggcaccgc aacaccagcg tctccatggt ggacttcagc	3120
atcgcagtgg agaatcagtt gtctggaaac ctgctgagga aattcaaaaa cagcaacggg	3180
tggcagaagc tgtgggtggt gtccacaaac ttctgcctgt tcttctacaa atcacaccag	3240
gacaatcatc cccttgccag cctgcctctg ctcggctact cgctcaccat cccctctgag	3300
tccgagaaca tccagaaaga ctacgtgttc aagctgcact tcaagtccca cgtctactac	3360
ttcagggcgg aaagcgagta cacgttcgaa aggtggatgg aagtgatccg cagtgccacc	3420
agctctgcct cgcgacccca cgtgttgagt cacaagagat ctcttggtga ttgatggccg	3480
gacacactcg tttccgcagt ggctgctttc ctggaagacg tttcctttct tctgtattaa	3540
tgaagcctgg taaaattaac acctgtctga aaatcaaaaa catggcttcc cagcagctct	3600
cctgtctcca cagccgcgtt ttttaacccc gacctctcag cgtctgaatg aacagcgctc	3660
ccacctccag tcctggcatc cgctgggggc gctgttcttt agctagtgcc agtattaaaa	3720
cattgtcatt acgagagtgc caaatgacat cttccctcca ccctgcccct gaaaaacagt	3780
acacacacat ccgttcaaca caagacaggg caagtgtttt tcttcctaaa aaaagttctt	3840
tcttttatta ttttcaccta ttggctgctg cattttacga agtggacttc ccggtgtttg	3900
tttgtttggt tgcaatacac tcagtgcagc ctttaagcaaa tgagatcatt ttcagatttc	3960
attttttttt tcagtctttc tacttttgta ataataggaa gttagtagga ctcacttctc	4020
tgattaataa gcaatttgca gcacacagcg ttccactgcg gggtttcacg ctcacctgaa	4080
aacacctgtt cccaacctac ttcttggtgc aagttgacca aatcgtttta agtggtaact	4140
ctttccaacc gtagcagggg tgttttctgt taagcaaagc cgagatccag tgcaatacct	4200
ggactgtcac cgtcctgtga gtggtgtaca caatgggaag ataataagcc gtggtgtttt	4260
gctgtctgtc tgtgtcacaa gcatgaaaac ccgtgtgtca ttgatcagca ccatttgagg	4320
tatgttccgt gatgagcgtt tagtgagcct gctggctgca gagcactatg aaatcatggt	4380
acgtagtccc cggcacctgt cgttattcct atatcctcct gcaactgtgg tttgaaactg	4440
cgcattctct agtagtatat atcgtgcctg tcttcaaaaa catttccctt tttataactca	4500
ttccccccag gcatggggta gtgtcagtcg gactgcacag ggaacacggg ttccagtggc	4560
tttggccccct actcgggaaa cgtctgcctg ttctcgatgg tgatgggggt gctgccattc	4620
ccctgggtttt cctaagccct ttctaacgag agtctcaaac aagcggaggc gagggccaat	4680
tcaaccccat tctttccagc gccccgcacc atagcacctg cccacctgag aaccaggaac	4740
gcacctctc tgtggagctc tgactgggtg agctggaaac aaacagcaac ttgcaaacgg	4800

acgaagagcc tgccgtgtgt taatcatttg ccttacaaga tgtaccagac ggtttccagt 4860  
 actaacaaag ggaataaaaa tacctcacgc cacaatccag catattgatg ttttaaggca 4920  
 aaacaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4980  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 5028

<210> 10

<211> 7787

<212> DNA

<213> NM\_012334.1| Homo sapiens myosin X (MYO10), mRNA

<400> 10  
 gagacaaagg ctgccgtcgg gacgggagag ttagggactt gggtttgggc gaacaaaagg 60  
 tgagaaggac aagaaggagc cgggagatgg cagcagggga gccccgcggg cgcgcgctct 120  
 cgggagtggt gccgtgacac gcatggtttc cccggacccg cggcggcgct gacttccgcg 180  
 agtcggagcg gcactcggcg agtccgggac tgcgctggaa caatggataa cttcttcacc 240  
 gagggaacac gggctctggct gagagaaaat ggccagcatt ttccaagtac tgtaaattcc 300  
 tgtgcagaag gcatcgctgt cttccggaca gactatggtc aggtattcac ttacaagcag 360  
 agcacaatta cccaccagaa ggtgactgct atgcacccca cgaacgagga gggcgtggat 420  
 gacatggcgt ccttgacaga gctccatggc ggctccatca tgtataactt attccagcgg 480  
 tataagagaa atcaaatata tacctacatc ggctccatcc tggcctccgt gaaccctac 540  
 cagcccatcg ccgggctgta cgagcctgcc accatggagc agtacagccg gcgccacctg 600  
 ggcgagctgc ccccgacat cttcgccatc gccaacgagt gctaccgctg cctgtggaag 660  
 cgctacgaca accagtgcac cctcatcagt ggtgaaagtg gggcaggtaa aaccgaaagc 720  
 actaaattga tcctcaagtt tctgtcagtc atcagtcaac agtctttgga attgtcctta 780  
 aaggagaaga catcctgtgt tgaacgagct attcttgaaa gcagcccat catggaagct 840  
 ttcggcaatg cgaagaccgt gtacaacaac aactctagtc gctttgggaa gtttgttcag 900  
 ctgaacatct gtcagaaagg aaatattcag ggcgggagaa ttgtagatta tttattagaa 960  
 aaaaaccgag tagtaaggca aaatcccggg gaaaggaatt atcacatatt ttatgcactg 1020  
 ctggcagggc tggaacatga agaaagagaa gaattttatt tatctacgcc agaaaactac 1080  
 cactacttga atcagtctgg atgtgtagaa gacaagacaa tcagtgacca ggaatccttt 1140  
 agggaagtta ttacggcaat ggacgtgatg cagttcagca aggaggaagt tcgggaagtg 1200  
 tcgaggctgc ttgctgggtat actgcatctt gggaacatag aatttatcac tgctgggtggg 1260  
 gcacagggtt ccttcaaaac agctttgggc agatctgcgg agttacttgg gctggaccca 1320  
 acacagctca cagatgcttt gaccagaga tcaatgttcc tcaggggaga agagatcctc 1380

acgcctctca	atgttcaaca	ggcagtagac	agcagggact	ccctggccat	ggctctgtat	1440
gcgtgctgct	ttgagtgggt	aatcaagaag	atcaacagca	ggatcaaagg	caatgaggac	1500
ttcaagtcta	ttggcatcct	cgacatcttt	ggatttgaaa	actttgaggt	taatcacttt	1560
gaacagttca	atataaacta	tgcaaacgag	aaacttcagg	agtacttcaa	caagcatatt	1620
ttttcttttag	aacaactaga	atatagccgg	gaaggattag	tgtgggaaga	tattgactgg	1680
atagacaatg	gagaatgcct	ggacttgatt	gagaagaaac	ttggcctcct	agcccttata	1740
aatgaagaaa	gccattttcc	tcaagccaca	gacagcacct	tattggagaa	gctacacagt	1800
cagcatgcga	ataaccactt	ttatgtgaag	cccagagttg	cagttaacaa	ttttggagtg	1860
aagcactatg	ctggagaggt	gcaatatgat	gtccgaggta	tcttggagaa	gaacagagat	1920
acatttcgag	atgaccttct	caatttgcta	agagaaagcc	gatttgactt	tatctacgat	1980
ctttttgaac	atgtttcaag	ccgcaacaac	caggatacct	tgaaatgtgg	aagcaaacat	2040
cggcggccta	cagtcagctc	acagttcaag	gactcactgc	attccttaat	ggcaacgcta	2100
agctcctcta	atcctttctt	tgttcgctgt	atcaagccaa	acatgcagaa	gatgccagac	2160
cagtttgacc	aggcggttgt	gctgaaccag	ctgcggtact	cagggatgct	ggagactgtg	2220
agaatccgca	aagctgggta	tgcggtccga	agaccctttc	aggactttta	caaaaggtat	2280
aaagtgctga	tgaggaatct	ggctctgcct	gaggacgtcc	gagggaaagt	cacgagcctg	2340
ctgcagctct	atgatgcctc	caacagcgag	tggcagctgg	ggaagaccaa	ggctcttctt	2400
cgagaatcct	tggaacagaa	actggagaag	cggaggggaag	aggaagtga	ccacgcggcc	2460
atggtgattc	gggcccattg	cttgggcctt	ttagcacgaa	aacaatacag	aaaggtcctt	2520
tattgtgtgg	tgataataca	gaagaattac	agagcattcc	ttctgaggag	gagatttttg	2580
cacctgaaaa	aggcagccat	agttttccag	aagcaactca	gaggtcagat	tgctcggaga	2640
gtttacagac	aattgctggc	agagaaaagg	gagcaagaag	aaaagaagaa	acaggaagag	2700
gaagaaaaga	agaaacggga	ggaagaagaa	agagaaagag	agagagagcg	aagagaagcc	2760
gagctccgcg	cccagcagga	agaagaaacg	aggaagcagc	aagaactcga	agccttgtag	2820
aagagccaga	aggaagctga	actgaccctg	gaactggaga	aacagaagga	aaataagcag	2880
gtggaagaga	tcctccgtct	ggagaaagaa	atcgaggacc	tgacgcgcac	gaaggagcag	2940
caggagctgt	cgctgaccga	ggcttccctg	cagaagctgc	aggagcggcg	ggaccaggag	3000
ctccgcaggc	tggaggagga	agcgtgcagg	gcggcccagg	agttcctcga	gtccctcaat	3060
ttcgacgaga	tcgacgagtg	tgtccggaat	atcgagcggg	ccctgtcggg	gggaagcgaa	3120
ttttccagcg	agctggctga	gagcgcatgc	gaggagaagc	ccaacttcaa	cttcagccag	3180
ccctaccag	aggaggaggt	cgatgagggc	ttcgaagccg	acgacgacgc	cttcaaggac	3240
tcccccaacc	ccagcgagca	cggccactca	gaccagcgaa	caagtggcat	ccggaccagc	3300
gatgactctt	cagaggagga	cccatacatg	aacgacacgg	tgggtccccc	cagccccagt	3360
gcggacagca	cggtgctgct	cgccccatca	gtgcaggact	ccgggagcct	acacaactcc	3420

tccagcggcg agtccaccta ctgcatgccc cagaacgctg gggacttgcc ctccccagac	3480
ggcgactacg actacgacca ggatgactat gaggacggtg ccatcacttc cggcagcagc	3540
gtgaccttct ccaactccta cggcagccag tgggtccccg actaccgctg ctctgtgggg	3600
acctacaaca gctcgggtgc ctaccggttc agctctgagg gggcgcagtc ctcgtttgaa	3660
gatagtgaag aggactttga ttccagggtt gatacagatg atgagctttc ataccggcgt	3720
gactctgtgt acagctgtgt cactctgccg tatttccaca gctttctgta catgaaaggt	3780
ggcctgatga actcttgga acgccgctgg tgcgtcctca aggatgaaac cttcttggg	3840
ttccgctcca agcaggaggc cctcaagcaa ggctggctcc acaaaaaagg ggggggctcc	3900
tccacgctgt ccaggagaaa ttggaagaag cgctggtttg tcctccgcca gtccaagctg	3960
atgtactttg aaaacgacag cgaggagaag ctcaaggga ccgtagaagt gcgaacggca	4020
aaagagatca tagataacac caccaaggag aatgggatcg acatcattat ggccgatagg	4080
actttccacc tgattgcaga gtccccagaa gatgccagcc agtgggtcag cgtgctgagt	4140
caggctccacg cgtccacgga ccaggagatc caggagatgc atgatgagca ggcaaacc	4200
cagaatgctg tgggcacctt ggatgtgggg ctgattgatt ctgtgtgtgc ctctgacagc	4260
cctgatagac ccaactcggt tgtgatcatc acggccaacc ggggtgctgca ctgcaacgcc	4320
gacacgccgg aggagatgca ccactggata accctgctgc agagggtccaa aggggacacc	4380
agagtggagg gccaggaatt catcgtgaga ggatgggtgc acaaagagg gaagaacagt	4440
ccgaagatgt cttcactgaa actgaagaaa cggtgggttg tactcaccca caattccctg	4500
gattactaca agagttcaga gaagaacgcg ctcaaactgg ggaccctggt cctcaacagc	4560
ctctgctctg tcgtcccccc agatgagaag atattcaaag agacaggcta ctggaacgtc	4620
accgtgtacg ggcgcaagca ctgttaccgg ctctacacca agctgctcaa cgaggccacc	4680
cggtgggtcca gtgccattca aaacgtgact gacaccaagg ccccgatcga cccccacc	4740
cagcagctga ttcaagatat caaggagaac tgcctgaact cggatgtggt ggaacagatt	4800
tacaagcggg acccgatcct tcgatacacc catcacccct tgcactcccc gtcctgccc	4860
cttccgtatg gggacataaa tctcaacttg ctcaaagaca aaggctatac cacccttcag	4920
gatgaggcca tcaagatatt caattccctg cagcaactgg agtccatgtc tgaccaatt	4980
ccaataatcc agggcatcct acagacagg catgacctgc gacctctgcg ggacgagctg	5040
tactgccagc ttatcaaaca gaccaaaaa gtgccccacc ccggcagtggt gggcaacctg	5100
tacagctggc agatcctgac atgcctgagc tgcaccttcc tgccgagtcg agggattctc	5160
aagtatctca agttccatct gaaaaggata cgggaacagt ttccaggaac cgagatggaa	5220
aaatacgtc tcttactta cgaatctctt aagaaaacca aatgccgaga gtttgtgcct	5280
tcccagatg aaatagaagc tctgatccac aggcaggaaa tgacatccac ggtctattgc	5340
catggcggcg gctcctgcaa gatcaccatc aactcccaca ccactgctgg ggagggtggtg	5400

gagaagctga tccgaggcct ggccatggag gacagcagga acatgtttgc tttgtttgaa	5460
tacaacggcc acgtcgacaa agccattgaa agtcgaaccg tcgtagctga tgtcttagcc	5520
aagtttgaaa agctggctgc cacatccgag gttggggacc tgccatggaa attctacttc	5580
aaactttact gcttcctgga cacagacaac gtgccaaaag acagtgtgga gtttgcattt	5640
atgtttgaac aggcccacga agcggttatc catggccacc atccagcccc ggaagaaaac	5700
ctccagggtc ttgctgccct gcgactccag tatctgcagg gggattatac tctgcacgct	5760
gccatcccac ctctcgaaga ggtttattcc ctgcagagac tcaaggcccc catcagccag	5820
tcaacaaaaa ctttcacccc ttgtgaacgg ctggagaaga ggcggacgag cttcctagag	5880
gggaccctga ggcggagctt ccggacagga tccgtggtcc ggcagaaggt cgaggaggag	5940
cagatgctgg acatgtggat taaggaagaa gtctcctctg ctcgagccag tatcattgac	6000
aagtggagga aatttcaggg aatgaaccag gaacaggcca tggccaagta catggccttg	6060
atcaaggagt ggcctggcta tggctcgacg ctgtttgatg tggagtgcaa ggaagggtggc	6120
ttccctcagg aactctggtt ggggtgtcagc gcggacgccg tctccgtcta caagcgtgga	6180
gaggaagac cactggaagt cttccagtat gaacacatcc tctcttttgg ggcacccctg	6240
gcgaatacgt ataagatcgt ggtcgatgag agggagctgc tctttgaaac cagtgaggtg	6300
gtggatgtgg ccaagctcat gaaagcctac atcagcatga tcgtgaagaa gcgctacagc	6360
acgacacgct ccgccagcag ccagggcagc tccaggtgaa ggcgggacag agcccacctg	6420
tctttgctac ctgaacgcac caccctctgg cctaggctgg ctccagtgtg ccatgcccag	6480
ccaaaacaaa cacagagctg cccaggcttt ctggaagctt ctggtctgag ggaggtgtct	6540
ccgaggatcc ttttgccctgc cgccttcatt gatcctgtat taagctgtca actttaacag	6600
tctgcacagt ttccaaagct ttactactct tagaggacac atgccttaaa aaaggagggg	6660
aggaaccacg ctgccaccaa agcagccgga agtgccctaa cttgtggaac caacactaat	6720
cgaccgtaac tgtgctactg aagggaactg cttttcccc ttctggggga gacttaacag	6780
agcgtggaag gggggcattc tctgtcaatg atgcactaac ctcccaacct gatttccccg	6840
aatctgaggg aagggtgaggg agtggggaagg gggatggaga gctcgagggg acagtgtggt	6900
tgagctggag tgctgcgggc agcctttctc atggaatgac atgaatcaac tttttcttt	6960
gtttcatctt ttaagtgtac gtgcttgctt gttcgtgcat gtgttcataa actcaacact	7020
ttaatcatgg tttcatgagc attaaaaagc aaagggaaaa aggatgtgta atgggtgtaca	7080
cagtctgtat attttaataa tgcagagcta tagtctcaat tgttacttta taagggtggt	7140
ttattaacaa acccaaatcc tggattttcc tgtctttgct gtattttgaa aaacacgtgt	7200
tgactccatt gttttacatg tagcaaagtc tgccatctgt gtctgctgta ttataaacag	7260
ataagcagcc tacaagataa ctgtatttat aaaccactct tcaacagctg gctccagtgc	7320
tggttttaga acaagaatga agtcattttg gagtctttca tgtctaaaag atttaagtta	7380
aaaacaaagt gttacttgga aggttagctt ctatcattct ggatagatta cagatataat	7440

aaccatgttg actatggggg agagacgctg cattccagaa acgtcttaac acttgagtga	7500
atcttcaaag gaccctgaca ttaaatgctg aggcctttaat acacacatat tttatcccaa	7560
gtttataatg gtgggtctgaa caaggcacct gtaaataaat cagcatttat gaccagaaga	7620
aaaataatct ggtcttggac tttttatatt tatatggaaa agttttaagg acttgggcca	7680
actaagtcta cccacacgaa aaaagaaatt tgccttgccc ctttgtgtac aaccatgcaa	7740
aactgtttgt tggctcacag aagttctgac aataaaagat actagct	7787

<210> 11

<211> 2033

<212> DNA

<213> NM\_001533.1| Homo sapiens heterogeneous nuclear ribonucleoprotein L (HNRPL), mRNA

<400> 11	
ggacgagcag cggaggcggg cgggagcgat ggtgaagatg gcggcggcgg gcggcggagg	60
cggcgggtggc cgctactacg gcggcggcag tgagggcggc cggggccccta agcgggtcaa	120
gactgacaac gccggcgacc agcacggagg cggcggcggg ggcgggtggag gagccggggc	180
ggcggggcggc ggcgggcggg gggagaacta cgatgacccg caaaaaacc ctgcctcccc	240
agttgtccac atcagggggc tgattgacgg tgtgggtggaa gcagaccttg tggaggcctt	300
gcaggagttt ggacccatca gctatgtggt ggtaatgcct aaaaagagac aagcactggt	360
ggagtttgaa gatgtgttgg gggcttgcaa cgcagtgaac tacgcagccg acaaccaaat	420
atacattgct ggtcaccag cttttgtcaa ctactctacc agccagaaga tctcccgccc	480
tggggactcg gatgactccc ggagcgtgaa cagtgtgctt ctctttacca tcctgaaccc	540
catttattcg atcaccacgg atgttcttta cactatctgt aatccttggt gccctgtcca	600
gagaattgtc attttcagga agaattggagt tcaggcgatg gtggaatttg actcagttca	660
aagtgccag cgggccaagg cctctctcaa tggggctgat atctattctg gctgttgcac	720
tctgaagatc gaatacgcaa agcctacacg cttgaatgtg ttcaagaatg atcaggatac	780
ttgggactac acaaaccaca atctcagtgg acaagggtgac cctggcagca accccaacaa	840
acgccagagg cagccccctc tcctgggaga tcaccccgca gaatatggag ggccccacgg	900
tgggtaccac agccattacc atgatgaggg ctacggggccc cccccacctc actacgaagg	960
gagaaggatg ggtccaccag tgggggggtca ccgtcggggc ccaagtcgct acggccccca	1020
gtatgggcac cccccacccc ctccccacc acccgagtat ggccctcacg ccgacagccc	1080
tgtgctcatg gtctatggct tggatcaatc taagatgaac ggtgaccgag tcttcaatgt	1140
cttctgctta tatggcaatg tggagaagg gaaattcatg aaaagcaagc cggggggccgc	1200
catggtggag atggctgatg gctacgctgt agaccggggc attacccacc tcaacaacaa	1260

cttcatgttt	gggcagaagc	tgaatgtctg	tgtctccaag	cagccagcca	tcatgcctgg	1320
tcagtcatac	gggttggaag	acgggtcttg	cagttacaaa	gacttcagtg	aatcccggaa	1380
caatcggttc	tccacccag	agcaggcagc	caagaaccgc	atccagcacc	ccagcaacgt	1440
gctgcacttc	ttcaacgccc	cgctggaggt	gaccgaggag	aacttctttg	agatctgcga	1500
tgagctggga	gtgaagcggc	catcttctgt	gaaagtattc	tcaggcaaaa	gtgagcgag	1560
ctcctctgga	ctgctggagt	gggaatccaa	gagcgatgcc	ctggagactc	tgggcttcct	1620
gaaccattac	cagatgaaaa	acccaaatgg	tccataccct	tacactctga	agttgtgttt	1680
ctccactgct	cagcacgcct	cctaattagg	tgccataggaa	gagtcccatc	tgagcaggaa	1740
gacattttctc	tttcctttat	gccatttttt	gtttttgtta	tttgcaaaag	atcttgatt	1800
cctttttttt	ttttttttt	tttaaagtct	aggttttag	aggcttactt	aaccttaatg	1860
gaaacgctgg	aaatctgcag	ggggagggag	aggggaactg	ttatctccca	agattaacct	1920
tcacttttaa	aaaattattg	tacatgtgat	ttttttttt	cctgttcata	catttggtgt	1980
gcccattgtac	tcttggcaca	tttcaataaa	attgtttgga	aaataaacac	agc	2033

<210> 12

<211> 3453

<212> DNA

<213> NM\_001144.3| Homo sapiens autocrine motility factor receptor (AMFR), transcript variant 1, mRNA

<400> 12						60
gggccgccgc	agaggccccg	ccgcagcgca	gggaagcctg	ggggccagag	gtcgccgctg	
ccgccatgcc	gctgctcttc	ctcgagcgct	tcccctggcc	cagcctccgc	acctacacgg	120
gcctcagcgg	cctggccctg	ctgggcacca	tcatcagcgc	ctaccgcgcg	ctcagccagc	180
ccgaggccgg	ccccggcgag	ccggaccagc	taacggcctc	gctgcagcct	gagccgccgg	240
cgcccccccg	gccgagcgcc	gggggacccc	gggcccgcga	tgtggcccag	tacctgctct	300
cagacagcct	cttcgtgtgg	gttctagtaa	ataccgcttg	ctgtgttttg	atgttggtgg	360
ctaagctcat	ccagtgtatt	gtgtttggcc	ctcttcgagt	gagtgcagaga	cagcatctca	420
aagacaaatt	ttggaatttt	atcttctaca	agttcatttt	catctttggt	gtgctgaatg	480
tccagacagt	ggaagaggtg	gtcatgtggt	gcctctgggt	tgccggactt	gtctttctgc	540
acctgatggt	tcagctctgc	aaggatcgat	ttgaatatct	ttccttctcg	cccaccacgc	600
cgatgagcag	ccacggctga	gtcctgtccc	tgttgggtgc	catgctgctt	tcctgctgtg	660
gactggcggc	cgtctgctcc	atcaccggct	acaccacagg	aatgcacacc	ttggctttca	720
tggtcgcaga	gtctcttctt	gtgacagtga	ggactgctca	tgtgatttta	cgatacgtaa	780
ttcacctctg	ggacctcaac	cacgaaggga	cgtgggaagg	aaaggggacg	tatgtctatt	840



acacagactt tgtcatggag ctcactctcc tgtccctgga cctcatgcac catattcaca	900
tgttgttatt tggcaacatc tggttatcca tggccagcct ggtcatcttt atgcagctgc	960
gttacctgtt tcatgagggtg caacgtcgaa ttcgtcggca caagaactat ctacgtgtgg	1020
ttggaacat ggaggccagg tttgcagttg caactccaga ggagctggct gtcaacaatg	1080
acgactgtgc catctgttgg gactccatgc aggtgcgcg gaaactgccc tgtggacatc	1140
ttttccacaa ctctgtctt cgttcctggc tagaacaaga cacctcctgt ccaacatgca	1200
gaatgtctct taatattgcc gacaataatc gtgtcaggga agaacatcaa ggagagaact	1260
tggatgagaa tttggttcct gtagcagcag ccgaagggag acctcgctta aaccaacaca	1320
atcacttctt ccatttcgat gggctcggga ttgcgagctg gctgccgagt ttttcggttg	1380
aagtgatgca caccaccaac attcttggca ttacgcaggc cagcaactcc cagctcaatg	1440
caatggctca tcagattcaa gagatgtttc cccaggttcc ataccatctg gtactgcagg	1500
acctccagct gacacgctca gttgaaataa caacagacaa tatttttagaa ggacggattc	1560
aagtaccttt tcctacacag cggtcagata gcatcagacc tgcattgaac agtcctgtgg	1620
aaaggccaag cagtgaccag gaagagggag aaacttctgc tcagaccgag cgtgtgccac	1680
tggacctcag tcctcgcctg gaggagacgc tggacttcgg cgagggtggaa gtggagccca	1740
gtgaggtgga agacttcgag gctcgtggga gccgcttctc caagtctgct gatgagagac	1800
agcgcagtgt ggtgcagcgt aaggacgaac tcctccagca agctcgcaaa cgtttcttga	1860
acaaaagttc tgaagatgat gcggcctcag agagcttctt cccctcggaa ggtgcgtcct	1920
ctgaccccggt gacctgcgt cgaaggatgc tggctgccgc cgcggaacgg aggcttcaga	1980
agcagcagac ctctagcgc tcccttgctt tcctcagctg cctcctgcgc cctgtgcccc	2040
actgactgga ggaggcctgt cccaattctg cccgctccat ggaaaagcgg gcttgactgc	2100
attgccgctg tataaagcat gtggtcttat agtgtttgga cagctgataa atttaatcct	2160
tctttgtaat actttctatg tgacatttct ctccccctta gaaacactgc aaattttaac	2220
tgtaggatat atctcttctg gtgttgactg gactgcttgg ggtgggggac gatcaggagg	2280
aagtgagcag tcgcctgcct gcagcaggca gcttctactc ctgcctcatg catacgtccc	2340
acaaatgcag gtgtcctgag caccacaccc agtggggaaga gtgtggggga ggcgcacagt	2400
gtgagcccg cccacgctc tggggtaaca tctgttatca aactgctgtc gttgttgtgg	2460
aagcatgtag actgtgccag aggccagacc cacgggctca tgcacccctg agccagcagg	2520
gcatcttggga aaaggaactc ttggttcgat acctggagca gaggagggga aagtccaggg	2580
ctataggggtg tgatgaagtc acccctttct gtcccactac atctgggact gactttccga	2640
gcctccagtc caaagccggc ttgatttccg tgaactctgg tgctcctgca tctcatgagt	2700
gtgccccatg ggtccctcc cctctcagca tttccttgtc ccgtctggac ctggggagtg	2760
gttaggcagc aagctttggt ttatggtttt cattcattgg tgaagtaaat taggcagtgc	2820

taaagcctgt gggtttggc cttgaacaag atgtgggcct tgcaagatgg gagagtaaac	2880
cttgaagggc tttattaaag aaataaaaaa gaacttttgt atcttttatc ctgggagcac	2940
tgcgttttcc tagctgtgtt attcctgggt taattcagca gagaaggtaa ggtgtgaacc	3000
tacctgcctt ggagagggcc cagggtccca atctcttcaa attcttcaca tgtttaactt	3060
taaggatttg aaccatgaag tcataggtta cagacctcag ttttatgccc cattggatta	3120
cttttttttt tttttttttt tttttttttt tactcttga aagctttgtt ttgtggtagt	3180
ccttttggga agaatccagt attatctaca attattggca aagtttaaata gtattttaca	3240
taacggaaag tttttagaat gttgaaaagt aattgaaaaa ggtgatagggt aaatttttag	3300
gcaaagataa tttatttcaa taaatctttc aaaagcctta ccttgaaatg ctgttagtaa	3360
atttctgtga tttttttttt taatttgttt tgctgagagc atagctatatt gtttttattg	3420
taaaacaata ataataataa aaagcaaact cta	3453

<210> 13

<211> 1351

<212> DNA

<213> NM\_013974.1| Homo sapiens dimethylarginine dimethylaminohydrolase 2 (DDAH2), mRNA

<400> 13	
ccgcttagac aatgccccgg agccgccaga ccgtcgcgcc cctgccccat cgtagtatat	60
gagctcgctt acacaaggac ccccgctaaa agccagagct cccagtcccc gaggcttgaa	120
gacggggact cccttctcca ccaactctgt cctcgggggg tggggcccca gccgagatca	180
cagcgcgaca ggagtggggg tggccgctgg agacaggtga agaaacaaga aaactaagaa	240
atccgagcgg ttggaggggg agtctgtgtg gatgggatgg ggacgccggg ggaggggctg	300
ggccgctgct cccatgccct gatccgggga gtcccagaga gcctggcgctc gggggaagggt	360
gcgggggctg gccttcccgc tctggatctg gccaaagctc aaagggagca cggggtgctg	420
ggaggtaaac tgaggcaacg actggggcta cagctgctag aactgccacc tgaggagtca	480
ttgccgctgg gaccgctgct tggcgacacg gccgtgatcc aaggggacac ggccctaatac	540
acgcggccct ggagccccgc tcgtaggcca gaggtcgatg gagtccgcaa agccctgcaa	600
gacctggggc tccgaattgt ggaaatagga gacgagaacg cgacgctgga tggcactgac	660
gttctcttca ccggccggga gtttttcgta ggcctctcca aatggaccaa tcaccgagga	720
gctgagatcg tggcggacac gttccgggac ttcgccgtct cactgtgcc agtctcgggt	780
ccctcccacc tgcgcggtct ctgcggcatg gggggacctc gactgttgtt ggcaggcagc	840
agcgacgctg cccaaaaggc tgtccgggca atggcagtc tgacagatca cccatatgcc	900
tccctgaccc tcccagatga cgcagctgct gactgcctct ttcttcgtcc tgggttgctt	960

```

ggtgtgcccc ctttctcct gcaccgtgga ggtggggatc tgcccaacag ccaggaggca 1020
ctgcagaagc tctctgatgt caccctggta cctgtgtcct gctcagaact ggagaaggct 1080
ggcgccgggc tcagctccct ctgcttggtg ctcagcacac gccccacag ctgagggcct 1140
ggccttgggg tactgctggc caggggtagg atagtatagg aagtagaagg ggaaggaggg 1200
ttagatagag aatgctgaat aggcagtagt tgggagagag cctcaatatt gggggagggg 1260
agagtgtagg gaaaaggatc cactgggtga atcctccctc tcagaaccaa taaaatagaa 1320
ttgacctttt aaaaaaaaaa aaaaaaaaaa a 1351

```

<210> 14

<211> 4180

<212> DNA

<213> NM\_006291.2| Homo sapiens tumor necrosis factor, alpha-induced protein 2 (TNFAIP2), mRNA

```

<400> 14
ccagggtgat gctgaagatg atgaccttct tccaaggcct ctagagccat cagcctgtgc 60
caggcaccct cgacttgctt agaggccccc aaaagttgca gtccacatca gaggcagagt 120
cagaggcctc catgtcggag gcctcctctg aggacctggt gccaccctg gaggctgggg 180
cagccccata tagggaggag gaagaggcgg cgaagaagaa gaaggagaag aagaagaagt 240
ccaaaggcct ggccaatgtg ttctgctgtc tcaccaaagg gaagaagaag aagggtcagc 300
ccagctcagc ggagcccagag gacgcagccg ggtccaggca ggggctggat ggcccgcctc 360
ccacagtgga ggagctgaag gcggcgctgg agcgcgggca gctggaggcg gcgcggccgc 420
tgctggcgct ggagcgggag ctggcgggcg cggcggcggc gggcggtgtg agcgaggagg 480
agctggtgcg gcgccagagc aagggtggagg cgctgtacga gctgctgcgc gaccaggtgc 540
tgggcgtgct gcggcgggcg ctggaggcgc cgcccagagc gctgcgccag gcgctggccg 600
tgggtggcga gcaggagcgc gaggaccgcc aggcggcggc ggcggggccc gggacctcgg 660
ggctggcggc cagcgccccg cggcgctggc tgcagctgtg gcggcgcggc gtggcgaggg 720
cggccgagga gcgcatgggc cagcgggccg ccgcggggcg cgagggtccc gagagcgtct 780
ttctgcactt gggccgcacc atgaaggagg acctggaggc cgtggtggag cggctgaagc 840
cgctgttccc cgccgagttc ggcgtcgtgg cggcctacgc cgagagctac caccagcact 900
tcgcgggcca cctggccgcc gtggcgagc tgcagctgtg cgagcgcgac acctacatgc 960
tgctgctctg ggtgcagaac ctctacccca atgacatcat caacagcccc aagctgggtg 1020
gtgagctgca gggatatggg ctcgggagcc tcctgcccc caggcagatc cgactgctgg 1080
aggccacatt cctgtccagt gaggcggcca atgtgaggga gttgatggac cgagctctgg 1140
agctagaggc acggcgctgg gctgaggatg tgcctcccca gaggctggac ggccactgcc 1200

```

acagcgagct ggccatcgac atcatccaga tcacctccca ggcccaggcc aaggccgaga	1260
gcatcacgct ggacttgggc tcacagataa agcgggtgct gctgggtggag ctgcctgcgt	1320
tcctgaggag ctaccagcgc gcctttaatg aatttctgga gagaggcaag cagctgacga	1380
attacagggc caatgttatt gccaacatca acaactgcct gtccttccgg atgtccatgg	1440
agcagaattg gcaggtagcc caggacaccc tgagcctcct gctgggcccc ctgggtgagc	1500
tcaagagcca cggttttgac accctgctcc agaacctgca tgaggacctg aagccactgt	1560
tcaagagggt cacgcacacc cgctgggcgg cccctgtgga gacctggaa aacatcatcg	1620
ccactgtaga cagaggctg cctgagttct cagagctgca gggctgtttc cgggaggagc	1680
tcattggaggc cttgcacctg cacctggtga aggagtacat catccaactc agcaaggggc	1740
gcctggctct caagacggcc gagcagcagc agcagctggc tgggtacatc ctggccaatg	1800
ctgacaccat ccagcacttc tgcaccagc acggctcccc ggcgacctgg ctgcagcctg	1860
ctctccctac gctggccgag atcattcgcc tgcaggacct cagtgccatc aagattgagg	1920
tggccactta tgccacctgc taccctgact tcagcaaagg ccacctgagc gctatcctgg	1980
ccatcaaggg gaacctatcc aacagtgagg tcaagcgcat ccggagcatc ttggacgtca	2040
gcatgggggc gcaggagccc tcccggcccc tattttccct tataaagggt ggtagcttt	2100
tcctgtggcc tgacctgcct gtgagtggcc agcaagcctt gggcacaccc cgctgggagc	2160
tgtaagagc agcgtggtt ctcggttctt cccgggtctc ctgtgctctg atgctacttc	2220
tgcctagccc tggcggagggt gcaggccctg tcagctggaa ctggacagac cttggtttgt	2280
ttacatgtcc gatgggggca ggagctccca tcctgggcag ccaaccaggc aacaccaagg	2340
actctttgta aacgatagct gatcgtgtgc acgcaaggaa agaaccagga gggagagtgc	2400
agccaggctc agggatcccc ggacacctct gtccagagcc cctccacagt cggcctcatg	2460
actgtcctcc tcgtgggtgg ggccgagggc cctcttcagc tctctggaga caggggccga	2520
gcctcaccca tctgccctct gcagcccagg gccgccgtga gcgggattca gcaatggtgg	2580
aatggaagac agaactggaa gagaaagaag gaaaagatga gctctcgtct ggcaggggct	2640
tttaggggtc tgtggcgagc tgtgagcacc gccagcatta gacgtcacat ccaggtaggc	2700
ccacggcccc tacaggctgg ccctgcaatg gggccctgag ccctccctct tcatcccca	2760
aggcctcaac tagagggtgg tcccccgagg gcttgggtgt tactaccgaa gggccaaga	2820
cctcctgggt cctctcaggc tcccccttcc ccaaggcagg gacaggccct gggggtgcca	2880
ccgtgggccc tgccacccag aagtctggct gaggtctggg caggggcagg gcaagcttga	2940
cctctcactg ttgacccttt gccctctgta tttgtttcct attgccgtga caggtttcca	3000
caaacttcgt ggatcaaaac gaggtcttcc agttctgcgg gtcagaaggc tgaccgggg	3060
ctcaaactct ggtgtcggca gtcctgcact cttctggag gctctagggg agaattcatt	3120
tctggccttt tcatttttag aggctgaccg taattcttga cttcaggctc ctccatcttc	3180
agagccagct gtgggtagtt gaatcttttt cccgtcacct cattgaggcc tccccctctc	3240

tgccctccctc caccactttt tttttttttt ttttgagaca gggctcttgct gtgttgccca	3300
ggctggagtg cagtggcctg gtcattggcat caaggctcac tgcagcctgg acctcctggg	3360
tcaagtgate ctcttgcttc agtccccctga gacaatcccc cacgcccagc tacatatattt	3420
ttgtggatac aggggtctcat tctgttgccct aggcctgtct ggaactcctg ggctcaaggg	3480
atcttgtagc cttagcctcc taaagtgtctg ggattatagg catgagtcac tgtacccggc	3540
ctgctctacc gcttttaagg acgcttatga tcacattgcg cctaccaga gaaccaggt	3600
cgtctttcta ttttcaggtc agctgattag ccacctagt tccatctgca actttagttc	3660
ccactggctg tgtaacctaa catagtcaca ggctctgggg actgtcacgt ggacatcttt	3720
gggaggccgt tattctgccc accgcaccct ccgttcatcc cctgccctgc cgggcacctc	3780
gctctacccc aggaaaatgt gagctcgttt tcctgctcgg catgtgctcc ccctaaggct	3840
ctgctctcc ctgggcctga aagttccttc tcagcctgag agggggccct tcggactcag	3900
gcatgactca gcccggtga tgcctctgca gtgctgagtc aggatttggg gccggctctc	3960
ttgggtccgt ccccttttcc caggtagctg cttacaaagc tgtggccagg aagtggccgg	4020
tataaaggat gcccaaggctc tttgtacgtg tgtaggagtt agcgtgtttg atattgttaa	4080
tataataata attatttttt agagtactgc tttgtatgt atgttgaaca ggatccaggt	4140
ttttatagct tgatataaaa cagaattcaa aagtgaaaaa	4180

<210> 15

<211> 2524

<212> DNA

<213> NM\_000249.2| Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), mRNA

<400> 15	
attggctgaa ggcacttccg ttgagcatct agacgtttcc ttggctcttc tggcgccaaa	60
atgtcgttcg tggcaggggt tattcggcgg ctggacgaga cagtggtgaa ccgcatcgcg	120
gcgggggaag ttatccagcg gccagctaatt gctatcaaag agatgattga gaactgttta	180
gatgcaaaat ccacaagtat tcaagtgatt gttaaagagg gaggcctgaa gttgattcag	240
atccaagaca atggcaccgg gatcaggaaa gaagatctgg atattgtatg tgaaaggttc	300
actactagta aactgcagtc ctttgaggat ttagccagta tttctaccta tggctttcga	360
ggtgaggctt tggccagcat aagccatgtg gctcatgtta ctattacaac gaaaacagct	420
gatggaaagt gtgcatacag agcaagttac tcagatggaa aactgaaagc ccctcctaaa	480
ccatgtgctg gcaatcaagg gaccagatc acggtggagg acctttttta caacatagcc	540
acgaggagaa aagcttttaa aaatccaagt gaagaatatg ggaaaatttt ggaagttggt	600
ggcagggtatt cagtacacaa tgcaggcatt agtttctcag ttaaaaaaca aggagagaca	660

gtagctgatg ttaggacact acccaatgcc tcaaccgtgg acaatattcg ctccatcttt	720
ggaaatgctg ttagtcgaga actgatagaa attggatgtg aggataaaac cctagccttc	780
aaaatgaatg gttacatatc caatgcaaac tactcagtga agaagtgcac cttcttactc	840
ttcatcaacc atcgtctggt agaatcaact tccttgagaa aagccataga aacagtgtat	900
gcagcctatt tgcccaaaaa cacacacca ttctgtgacc tcagtttaga aatcagtcctc	960
cagaatgtgg atgttaatgt gcacccaca aagcatgaag ttcacttcct gcacgaggag	1020
agcatcctgg agcgggtgca gcagcacatc gagagcaagc tcctgggctc caattcctcc	1080
aggatgtact tcacccagac tttgctacca ggacttgctg gcccctctgg ggagatgggt	1140
aaatccacaa caagtctgac ctctcttct acttctggaa gtagtgataa ggtctatgcc	1200
caccagatgg ttcgtacaga ttcccgggaa cagaagcttg atgcatttct gcagcctctg	1260
agcaaaccctc tgtccagtca gcccaggcc attgtcacag aggataagac agatatttct	1320
agtggcaggg ctaggcagca agatgaggag atgcttgaac tcccagcccc tgctgaagtg	1380
gctgccaaaa atcagagctt ggagggggat acaacaaagg ggacttcaga aatgtcagag	1440
aagagaggac ctacttcag caacccaga aagagacatc gggaagattc tgatgtggaa	1500
atggtggaag atgattcccc aaaggaaatg actgcagctt gtaccccccg gagaaggatc	1560
attaacctca ctagtgtttt gagtctccag gaagaaatta atgagcaggg acatgagggt	1620
ctccgggaga tgttgcataa ccactccttc gtgggctgtg tgaatcctca gtgggccttg	1680
gcacagcatc aaaccaagtt ataccttctc aacaccacca agcttagtga agaactgttc	1740
taccagatac tcatttatga ttttgccaat tttggtgttc tcaggttatc ggagccagca	1800
ccgctctttg accttgccat gcttgcccta gatagtccag agagtggctg gacagaggaa	1860
gatggtccca aagaaggact tgctgaatac attgttgagt ttctgaagaa gaaggctgag	1920
atgcttgacag actatttctc tttggaaatt gatgaggaag ggaacctgat tggattaccc	1980
cttctgattg acaactatgt gcccccttg gagggactgc ctatcttcat tcttcgacta	2040
gccactgagg tgaattggga cgaagaaaag gaatgttttg aaagcctcag taaagaatgc	2100
gctatgttct attccatccg gaagcagtac atatctgagg agtcgaccct ctcaggccag	2160
cagagtgaag tgccctggctc cattccaaac tcctggaagt ggactgtgga acacattgtc	2220
tataaagcct tgcgtcaca cattctgcct cctaaacatt tcacagaaga tggaaatatc	2280
ctgcagcttg ctaacctgcc tgatctatac aaagtctttg agagggtgta aatatgggta	2340
tttatgcact gtgggatgtg ttcttctttc tctgtattcc gatacaaagt gttgtatcaa	2400
agtgtgatat acaagtgta ccaacataag tgttggtagc acttaagact tatacttgcc	2460
ttctgatagt attcctttat acacagtgga ttgattataa ataaatagat gtgtcttaac	2520
ataa	2524

<210> 16  
 <211> 1536  
 <212> DNA  
 <213> NM\_001071.1| Homo sapiens thymidylate synthetase (TYMS), mRNA

```

<400> 16
gggggggggg ggaccacttg gcctgcctcc gtcccgccgc gccacttggc ctgcctccgt    60
cccgccgcgc cacttcgcct gcctccgtcc cccgcccgcc gcgccatgcc tgtggccggc    120
tcggagctgc cgcgccggcc cttgcccccc gccgcacagg agcgggacgc cgagccgcgt    180
ccgccgcacg gggagctgca gtacctgggg cagatccaac acatcctccg ctgcggcgctc    240
aggaaggacg accgcacggg caccggcacc ctgtcgggtat tcggcatgca ggcgcgctac    300
agcctgagag atgaattccc tctgctgaca accaaacgtg tgttctggaa ggggtgttttg    360
gaggagttgc tgtggtttat caagggtacc acaaatgcta aagagctgtc ttccaagggg    420
gtgaaaatct gggatgccaa tggatcccga gacttttttg acagcctggg attctccacc    480
agagaagaag gggacttggg cccagtttat ggcttccagt ggaggcattt tggggcagaa    540
tacagagata tggaatcaga ttattcagga cagggagttg accaactgca aagagtgatt    600
gacaccatca aaaccaaccc tgacgacaga agaatcatca tgtgcgcttg gaatccaaga    660
gatcttcctc tgatggcgct gcctccatgc catgccctct gccagttcta tgtggtgaac    720
agtgagctgt cctgccagct gtaccagaga tcgggagaca tgggcctcgg tgtgcctttc    780
aacatcgcca gctacgccct gctcacgtac atgattgcgc acatcacggg cctgaagcca    840
ggtgacttta tacacacttt gggagatgca catatttacc tgaatcacat cgagccactg    900
aaaattcagc ttcagcgaga acccagacct ttcccaaagc tcaggattct tcgaaaagtt    960
gagaaaattg atgacttcaa agctgaagac tttcagattg aagggtacaa tccgcatcca   1020
actattaaaa tggaaatggc tgtttagggt gctttcaaag gagcttgaag gatattgtca   1080
gtcttttagg gttgggctgg atgccgaggt aaaagttctt tttgctctaa aagaaaaagg   1140
aactaggtca aaaatctgtc cgtgacctat cagttattaa tttttaagga tgttgccact   1200
ggcaaatgta actgtgccag ttctttccat aataaaaggc tttgagttaa ctactgagg   1260
gtatctgaca atgctgaggt tatgaacaaa gtgaggagaa tgaaatgtat gtgctcttag   1320
caaaaacatg tatgtgcatt tcaatcccac gtacttataa agaagggttg tgaatttcac   1380
aagctatttt tggaatatat ttagaatatt ttaagaattt cacaagctat tccctcaaat   1440
ctgagggagc tgagtaacac catcgatcat gatgtagagt gtggttatga actttatagt   1500
tgttttatat gttgctataa taaagaagtg ttctgc                                1536
  
```

<210> 17

<211> 2986

<212> DNA

<213> NM\_000201.1| Homo sapiens intercellular adhesion molecule 1 (CD54), human rhinovirus receptor (ICAM1), mRNA

<400> 17  
gcgccccagt cgacgctgag ctctctgct actcagagtt gcaacctcag cctcgctatg 60  
gtccccagca gccccggcc cgcgctgccc gcactcctgg tcctgctcgg ggctctgttc 120  
ccaggacctg gcaatgcca gacatctgtg tccccctcaa aagtcacct gccccgggga 180  
ggctccgtgc tgggtgacatg cagcacctcc tgtgaccagc ccaagttggt gggcatagag 240  
accccgttgc ctaaaaagga gttgctcctg cctgggaaca accggaaggt gtatgaactg 300  
agcaatgtgc aagaagatag ccaaccaatg tgctattcaa actgccctga tgggcagtca 360  
acagctaaaa ccttcctcac cgtgtactgg actccagaac gggtggaact ggcacccctc 420  
ccctcttggc agccagtggt caagaacctt accctacgct gccaggtgga ggggtgggca 480  
ccccgggcca acctcaccgt ggtgctgctc cgtggggaga aggagctgaa acgggagcca 540  
gctgtggggg agcccgctga ggtcacgacc acggtgctgg tgaggagaga tcaccatgga 600  
gccaatctct cgtgccgcac tgaactggac ctgcggcccc aagggctgga gctgtttgag 660  
aacacctcgg cccctacca gctccagacc tttgtcctgc cagcgactcc cccacaactt 720  
gtcagcccc gggctcctaga ggtggacacg caggggaccg tggctctgtc cctggacggg 780  
ctgttccag tctcggaggc ccaggtccac ctggcactgg gggaccagag gttgaacccc 840  
acagtcacct atggcaacga ctcttctcg gccaaaggct cagtcagtgt gaccgcagag 900  
gacgagggca ccagcggct gacgtgtgca gtaatactgg ggaaccagag ccaggagaca 960  
ctgcagacag tgaccatcta cagctttccg gcgccaacg tgattctgac gaagccagag 1020  
gtctcagaag ggaccgaggt gacagtgaag tgtgaggccc accctagagc caaggtgacg 1080  
ctgaatgggg ttccagccca gccactgggc ccgagggccc agctcctgct gaaggccacc 1140  
ccagaggaca acgggcgcag cttctcctgc tctgcaaccc tggaggtggc cggccagctt 1200  
atacacaaga accagacccg ggagcttcgt gtctgtatg gccccgact ggacgagagg 1260  
gattgtccgg gaaactggac gtggccagaa aattcccagc agactccaat gtgccaggct 1320  
tgggggaacc cattgcccga gctcaagtgt ctaaaggatg gcactttccc actgcccac 1380  
ggggaatcag tgactgtcac tcgagatctt gagggcacct acctctgtcg ggccaggagc 1440  
actcaagggg aggtcacccg cgaggtgacc gtgaatgtgc tctcccccg gtatgagatt 1500  
gtcatcatca ctgtggtagc agccgcagtc ataatgggca ctgcaggcct cagcacgtac 1560  
ctctataacc gccagcggaa gatcaagaaa tacagactac aacaggccca aaaagggacc 1620  
cccatgaaac cgaacacaca agccacgcct ccctgaacct atcccgggac agggcctctt 1680  
cctcggcctt cccatattgg tggcagtggt gccacactga acagagtgga agacatatgc 1740



catgcagcta cacctaccgg ccctgggacg ccggaggaca gggcattgtc ctcagtcaga	1800
tacaacagca tttggggcca tggtagctgc acacctaata cactaggcca cgcattctgat	1860
ctgtagtcac atgactaagc caagaggaag gagcaagact caagacatga ttgatggatg	1920
ttaaagtcta gcctgatgag aggggaagtg gtgggggaga catagcccca ccatgaggac	1980
atacaactgg gaaatactga aacttgctgc ctattgggta tgctgaggcc cacagactta	2040
cagaagaagt ggccctccat agacatgtgt agcatcaaaa cacaaggcc cactctcct	2100
gacggatgcc agcttgggca ctgctgtcta ctgacccaa cccttgatga tatgtattta	2160
ttcatttgtt attttaccag ctatttattg agtgtctttt atgtaggcta aatgaacata	2220
ggctcttggc ctcacggagc tcccagtcca tgtcacattc aaggtcacca ggtacagttg	2280
tacaggttgt acactgcagg agagtgcctg gcaaaaagat caaatggggc tgggacttct	2340
cattggccaa cctgcctttc ccagaagga gtgatttttc tatcggcaca aaagcactat	2400
atggactggg aatgggtcac aggttcagag attaccaggt gaggccttat tcctcccttc	2460
ccccaaaac tgacaccttt gttagccacc tccccacca catacatttc tgccagtgtt	2520
cacaatgaca ctacagcgtc atgtctggac atgagtgcc agggaatatg cccaagctat	2580
gccttgtcct cttgtcctgt ttgcatttca ctgggagctt gcactattgc agctccagtt	2640
tcctgcagtg atcaggggtc tgcaagcagt ggggaagggg gccaagggtat tggaggactc	2700
cctcccagct ttggaagggg catccgcgtg tgtgtgtgtg tgtatgtgta gacaagctct	2760
cgctctgtca ccagggctgg agtgcagtg tgcaatcatg gttcactgca gtcttgacct	2820
tttgggtcga agtgatcctc ccacctcagc ctctgagta gctgggacca taggctcaca	2880
acaccacacc tggcaaattt gatttttttt ttttttttca gagacggggg ctcgcaacat	2940
tgcccagact tcctttgtgt tagttaataa agctttctca actgcc	2986

<210> 18

<211> 736

<212> DNA

<213> NM\_004492.1| Homo sapiens general transcription factor IIA, 2 (12kD subunit) (GTF2A2), mRNA

<400> 18	
cgagctggag aggtggctcg agaagtagga acctcctgcc gggctcgtgg cggcttctgt	60
ccgctccgcg gaggaagcg ccttccccac aggacatcaa tgcaagcttg aataagaaaa	120
acaaattctt cctcctaagc catggcatat cagttataca gaaatactac tttgggaaac	180
agtcttcagg agagcctaga tgagctcata cagtctcaac agatcacccc ccaacttgcc	240
cttcaagttc tacttcagtt tgataaggct ataaatgcag cactggctca gagggtcagg	300
aacagagtca atttcagggg ctctctaaat acgtacagat tctgcgataa tgtgtggact	360

tttgtactga atgatgttga attcagagag gtgacagaac ttattaaagt ggataaagtg	420
aaaattgtag cctgtgatgg taaaaatact ggctccaata ctacagaatg aatagaaaaa	480
atatgacttt ttacaccat cttctgttat tcattgcttt tgaagagaag catagaagag	540
actttttatt tattctagaa ttgcagaaat gactacactg tgctatacca gagaattcca	600
gtagaaagaa acttgtaact ctgtagcctc ttacatcacc ttattatac agcatgaaaa	660
accataactt ttttttaagg acaaaagtgt ttgccttcct aagaaccttc ttttaataaac	720
tcattttaaa actctg	736

<210> 19

<211> 6401

<212> DNA

<213> NM\_004850.3| Homo sapiens Rho-associated, coiled-coil containing protein kinase 2 (ROCK2), mRNA

<400> 19	
caaggcggcc ggcggcgacc atggcagcgg gccggcgggc gccgtagtgg cccaggcctg	60
ggcttcagcc tcccggggcc ccagagggcg gggcggtccg ggccgcggcg gtggcggcg	120
cacttcctg ctcccggccg aggactcctg cgggcactcg ctgaggacca gcggaccggc	180
ggcggaatc tgactgaggg gcggggacgc cgtctgttcc ccgccgctcc cggcagggcc	240
gggcccgggct gggccgggct gggccgggcg ggcccctggg agcagccccc aggcggggga	300
ccgccttggg gacccgaagc cggagctaga ggcaggcggg gggcccgggt ggagtcgccg	360
ccggagctgg tggttcgggg gcggtgctag gccccgaggc tgcgggacct gagcgcgagg	420
agcctgagtg cgggtccagc ggtggcggca tgagccggcc cccgccgacg gggaaaatgc	480
ccggcgcccc cgagaccgcg ccgggggacg gggcaggcgc gagccgccag aggaagctgg	540
aggcgtgat ccgagacct cgctccccc tcaacgtgga gagcttgctg gatggcttaa	600
attccttggg ccttgattta gattttcctg ctttgaggaa aaacaagaac atagataatt	660
tcttaaatag atatgagaaa attgtgaaaa aaatcagagg tctacagatg aaggcagaag	720
actatgatgt tgtaaaagtt attggaagag gtgcttttgg tgaagtgcag ttggttcgtc	780
acaaggcatc gcagaagggt tatgctatga agcttccttag taagtttgaa atgataaaaa	840
gatcagattc tgcctttttt tgggaagaaa gagatattat ggcctttgcc aatagcccct	900
gggtggttca gcttttttat gcctttcaag atgatatgta tctgtacatg gtaatggagt	960
acatgcctgg tggagacct gtaaacctta tgagtaatta tgatgtgcct gaaaaatggg	1020
ccaaatttta cactgctgaa gttgttcttg ctctggatgc aatacactcc atgggtttta	1080
tacacagaga tgtgaagcct gacaacatgc tcttggataa acatggacat ctaaaattag	1140
cagattttgg cacgtgtatg aagatggatg aaacaggcat ggtacattgt gatacagcag	1200

ttggaacacc ggattatata tcacctgagg ttctgaaatc acaagggggt gatggtttct	1260
atgggcgaga atgtgattgg tggctctgtag gtgttttcct ttatgagatg ctagtggggg	1320
atactccatt ttatgcggat tcacttgtag gaacatatag caaaattatg gatcataaga	1380
attcactgtg tttccctgaa gatgcagaaa tttccaaaca tgcaaagaat ctcatctgtg	1440
ctttcttaac agatagggag gtacgacttg ggagaaatgg ggtggaagaa atcagacagc	1500
atcctttctt taagaatgat cagtggcatt gggataacat aagagaaacg gcagctcctg	1560
tagtacctga actcagcagt gacatagaca gcagcaattt cgatgacatt gaagatgaca	1620
aaggagatgt agaaaccttc ccaattccta aagcttttgt tggaaatcag ctgcctttca	1680
tcggatttac ctactataga gaaaatttat tattaagtga ctctccatct tgtagagaaa	1740
ctgattccat acaatcaagg aaaaatgaag aaagtcaaga gattcagaaa aaactgtata	1800
cattagaaga acatcttagc aatgagatgc aagccaaaga ggaactggaa cagaagtgca	1860
aatctgttaa tactcgcta gaaaaaacag caaaggagct agaagaggag attaccttac	1920
ggaaaagtgt ggaatcagca ttaagacagt tagaaagaga aaaggcgctt cttcagcaca	1980
aaaatgcaga atatcagagg aaagctgac atgaagcaga caaaaaacga aatttgga	2040
atgatgttaa cagcttaaaa gatcaacttg aagatttgaa aaaaagaaat caaaactctc	2100
aaatatccac tgagaaagt aatcaactcc agagacaact ggatgaaacc aatgctttac	2160
tgcgaaacaga gtctgatact gcagcccggg taaggaaaac ccaggcagaa agttcaaaac	2220
agattcagca gctggaatct aacaatagag atctacaaga taaaaactgc ctgctggaga	2280
ctgccaagtt aaaacttgaa aaggaattta tcaatcttca gtcagctcta gaatctgaaa	2340
ggagggatcg aacccatgga tcagagataa ttaatgattt acaaggtaga atatgtggcc	2400
tagaagaaga tttaaagaac ggcaaaatct tactagcgaa agtagaactg gagaagagac	2460
aacttcagga gagatttact gatttggaag aggaaaaaag caacatggaa atagatatga	2520
cataccaact aaaagttata cagcagagcc tagaacaaga agaagctgaa cataaggcca	2580
caaaggcacg actagcagat aaaaataaga tctatgagtc catcgaagaa gccaaatcag	2640
aagccatgaa agaaatggag aagaagctct tggaggaaag aactttaaaa cagaaagtgg	2700
agaacctatt gctagaagct gagaaaagat gttctctatt agactgtgac ctcaaacagt	2760
cacagcagaa aataaatgag ctctttaaac agaaagatgt gctaaatgag gatgttagaa	2820
acctgacatt aaaaatagag caagaaactc agaagcgctg ccttacacaa aatgacctga	2880
agatgcaaac acaacagggt aacacactaa aaatgtcaga aaagcagtta aagcaagaaa	2940
ataaccatct catggaaatg aaaatgaact tggaaaaaca aaatgctgaa cttcgaaaag	3000
aacgtcagga tgcagatggg caaatgaaag agctccagga tcagctcgaa gcagaacagt	3060
atttctcaac cttttataaa acacaagtta gggagcttaa agaagaatgt gaagaaaaga	3120
ccaaacttgg taaagaattg cagcagaaga aacaggaatt acaggatgaa cgggactctt	3180

tggtgcccc	actggagatc	accttgacca	aagcagattc	tgagcaactg	gctcgttcaa	3240
ttgctgaaga	acaatatctt	gatttgga	aagagaagat	catgaaagag	ctggagatca	3300
aagagatgat	ggctagacac	aaacaggaac	ttacggaaaa	agatgctaca	attgcttctc	3360
ttgaggaaac	taataggaca	ctaactagt	atgttgccaa	tcttgcaaat	gagaaagaag	3420
aattaaataa	caaattgaaa	gatgttcaag	agcaactgtc	aagattgaaa	gatgaagaaa	3480
taagcgcagc	agctattaaa	gcacagtttg	agaagcagct	attaacagaa	agaacactca	3540
aaactcaagc	tgtgaataag	ttggctgaga	tcatgaatcg	aaaagaacct	gtcaagcgtg	3600
gtaatgacac	agatgtgcgg	agaaaagaga	aggagaatag	aaagctacat	atggagctta	3660
aatctgaacg	tgagaaattg	accagcaga	tgatcaagta	tcagaaagaa	ctgaatgaaa	3720
tgcaggcaca	aatagctgaa	gagagccaga	ttcgaattga	actgcagatg	acattggaca	3780
gtaaagacag	tgacattgag	cagctgcggt	cacaactcca	agccttgc	attggtctgg	3840
atagttccag	tataggcagt	ggaccagggg	atgctgaggc	agatgatggg	tttcagaat	3900
caagattaga	aggatggctt	tcattgcctg	tacgaaacaa	cactaagaaa	tttgatggg	3960
ttaaaaagta	tgtgattgta	agcagtaaga	agattctttt	ctatgacagt	gaacaagata	4020
aagaacaatc	caatccttac	atggttttag	atatagacaa	gttatttcat	gtccgaccag	4080
ttacacagac	agatgtgtat	agagcagatg	ctaaagaaat	tccaaggata	ttccagattc	4140
tgtatgccaa	tgaaggagaa	agtaagaagg	aacaagaatt	tccagtggag	ccagttggag	4200
aaaaatctaa	ttatatattg	cacaagggac	atgagtttat	tcctactctt	tatcatttcc	4260
caaccaactg	tgaggcttgt	atgaagcccc	tgtggcacat	gtttaagcct	cctcctgctt	4320
tggagtgcg	ccgttgccat	attaagtgtc	ataaagatca	tatggacaaa	aaggaggaga	4380
ttatagcacc	ttgcaaagta	tattatgata	tttcaacggc	aaagaatctg	ttattactag	4440
caaattctac	agaagagcag	cagaagtggg	ttagtcgggt	ggtgaaaaag	atacctaaaa	4500
agcccccagc	tccagaccct	tttgcccgat	catctcctag	aacttcaatg	aagatacagc	4560
aaaaccagtc	tattagacgg	ccaagtcgac	agcttgcccc	aaacaaacct	agctaactgc	4620
cttctatgaa	agcagtcatt	attcaagggt	atcgtattct	tccagtgaaa	acaagactga	4680
aatatgatgg	cccaaaattt	attaaaaagc	tatatatttcc	tgagagactg	atacatacac	4740
tcatacatat	atgtgttccc	cttttccctg	taatataaat	tacaaatctg	ggctcctttg	4800
aagcaacagg	ttgaaccaac	aatgattggg	tgatagacta	aggatatatg	caactcttcc	4860
agacttttcc	ataaagctct	ctcggcagtc	gctcacacta	caatgcacac	aaggattgag	4920
aagagttaaa	ggctaaagaa	aacatctttt	ctagcttcaa	cagagagggt	tcaccagcac	4980
atttaccaga	agaatctggg	aatggattcc	actacagtga	tattgactgc	atctttaaga	5040
agtgaccatt	atactgtgta	tatatatata	aacacacaca	catatatata	tatatatata	5100
gtactcta	actgcaagaa	ggttttttta	acttcccact	ttatttttta	tacacattaa	5160
tcagatatca	ttacttgctg	cagttgcaac	tatgcacttg	tataaagcca	taatgttgga	5220

```

gtttatatca ctcatcctg tgtacctgat ggaagttgca tgttcatgtt taagcagtta 5280
ctgtaacaag aagtttaaag ttaattatat cagtttccta atgcttcatg ataggcaact 5340
ttacccattt tgaatgcctt aatttaattt ttttcaaagt ctgagccctg tctgtattaa 5400
aaaacaaaaa aagcgtttac cagctcttag gatgtaaact agctttgtgg aagataaatc 5460
gtgcactatt ttacacata aatagttata tcaatgtcag cctattttga ttaacaaatg 5520
tttttaaagt attattgggt atagaaacaa taatggatgg tgttggaact aatataatcct 5580
tgatgtctgt ctattattca ttcaactctt ttacagacc tcagtattag tctgtgacta 5640
caaaatatatt tatttgctt aaatttgctg gctaccctag atgtgttttt attcctggta 5700
aagacatttg tgattacatt ttcacactta agattcaaaa tttttcccaa atataaagaa 5760
aactaagaca gactgtagat gcattttaaa tatttaaata tgatcctcag acatgcagct 5820
gtgtgtggca gtattttagt accgggttaa gaaaactggc aactgggaag aagtggcctc 5880
aaaggcactt aatttgattt ttatttttta aatgctgtca aagttacagt ttacgcagga 5940
cattcttgcc gtattctcat gatcccagat aagtgtgtgt ttatactgc aacaatatgc 6000
agcaatggta agcgtaaagt tttttttttg tttttgtttt tttttatatt atgaagtctt 6060
ttaacagtct ctctttatat aaatacacag agtttggtat gatatttaaa tacatcatct 6120
ggccaggcat ggtggcttac gcctgtaatc ctagcacttt gggaggccaa gacgggcgga 6180
tcacctgagg tgaggagttc aagaccagcc tgcccaacat agtgaaactc cgtctctacc 6240
aatatacaaa aattagccgg gcattgatgg ggtggcctgt aatcccagct acttgggagg 6300
ctgagacagg agaatcgctt gaaccagga gacggtggtt gcagtgagcg aagatcgagc 6360
cactgcactc cagcctgggc agctgaacaa gactccgtct c 6401

```

<210> 20

<211> 1556

<212> DNA

<213> NM\_005783.3| Homo sapiens thioredoxin domain containing 9 (TXNDC9), mRNA

```

<400> 20
ggcgtccaag gtgatatcgc gcgaggttcg cagccaataa ggaggcggat gtgacggccc 60
gtttgcagcc gccggcagct actgcaaggc aaaagccgga gtggacgtgt cttttgaaac 120
tgctgctctt tcaattctca ggcgtcaccg agagctcagc acccaggctg aactctgtac 180
catttggaag aatggaagct gatgcatctg ttgacatgtt ttccaaagtc ctggagcatc 240
agctgcttca gactacaaa ctggtggaag aacatttgga ttctgaaatt caaaaactgg 300
atcagatgga tgaggatgaa ttggaacgcc ttaaagaaaa gagactccag gcactaagga 360
aagctcaaca gcagaaacaa gaatggcttt ctaaaggaca tggggaatac agagaaatcc 420

```

ctagtgaag agactttttt caagaagtca aggagagtga aaatgtggtt tgccatttct	480
acagagactc cacattcagg tgtaaaatac tagacagaca tctggcaata ttgtccaaga	540
aacacctcga gaccaaattt ttgaagctga atgtggaaaa agcacctttc ctttgtgaga	600
gactgcatat caaagtcatt cccacactag cactgctaaa agatgggaaa acacaagatt	660
atgttgttgg gtttactgac ctaggaaata cagatgactt caccacagaa actttagaat	720
ggaggctcgg ttcttctgac attcttaatt acagtggaaa tttaatggag ccaccatttc	780
agaacaaaaa gaaatttgga acaaaacttca caaagctgga aaagaaaact atccgaggaa	840
agaaatatga ttcagactct gatgatgatt agagctcaat aattctttgt aaattgtctt	900
tttttttctg cttcagattt aaatgtgttt ttaaaattct attaatgtct atacattggt	960
cacctaaata ctcatattct cgagttttat acagtgttat cacatcgaaa agtgtcttta	1020
ctgttttctg tgtggccatc atgtttaagt tgaggaaaac tcagttctta aattatctgg	1080
gaagggtctg gattctctat ttttgagatt gactttatca caatatgatt cttacatctt	1140
tataccattt acaatttgtt tttagatcta cagagttaga aattcgaaaa ctattccagg	1200
actaattctt aatcggcatt atttatacaa gaggtcaagt aacatttact agcgcaatac	1260
tgcacttgta aatgaattat aaacgctctt ctggaatata ttaaataaac cattaaagaa	1320
ctgcttattc attctggaca ctgcatgttg atgttgaatc aactgatgcc agcagaaagc	1380
tattttgatt tgtgaacata ctgccttatt taaagggtcc tgattgcttg tattttaaga	1440
cattcattaa aaagaaacca ggaaacactt ttgaaataac agcataagga acttcactgt	1500
ctctgctcaa taaaatacct gtaactggaa aaaaaaaaaa aaaaaaaaaa aaaaaa	1556

<210> 21

<211> 1276

<212> DNA

<213> NM\_003581.1| Homo sapiens NCK adaptor protein 2 (NCK2), mRNA

<400> 21	
gtgccaaaga aggactccat gaaagatgac agaagaagtt attgtgatag ccaagtggga	60
ctacaccgcc cagcaggacc aggagctgga catcaagaag gtgaacgagc ggctgtggtt	120
gctggacgac tccaagacgt ggtggcgggt gaggaacgcg gccaacagga cgggctatgt	180
accgtccaac tacgtggagc ggaagaacag cctgaagaag ggctccctcg tgaagaacct	240
gaaggacaca ctaggcctcg gcaagacgcg caggaagacc agcgcgcggg atgcgtcccc	300
cacgcccagc acggacgccc agtaccctgc caatggcagc ggcgccgacc gcatctacga	360
cctcaacatc ccggccttcg tcaagttcgc ctatgtggcc gagcgggagg atgagttgtc	420
cctggtgaag gggtcgcgcg tcaccgtcat ggagaagtgc agcgacggtt ggtggcgggg	480
cagctacaac gggcagatcg gctggttccc ctccaactac gtcttggagg aggtggacga	540

ggcggctgcg gagtcccca gcttcctgag cctgcgcaag ggcgcctcgc tgagcaatgg 600  
 ccaggggtcc cgcgtgctgc atgtgggtcca gacgctgtac cccttcagct cagtcaccga 660  
 ggaggagctc aacttcgaga aggggggagac catggagggtg attgagaagc cggagaacga 720  
 ccccgagtgg tggaaatgca aaaatgcccc gggccagggtg ggcctcgtcc ccaaaaacta 780  
 cgtggtgggtc ctcagtgcg ggcctgccct gcaccctgcg cacgccccac agataagcta 840  
 caccgggccc tcgtccagcg ggcgcttcgc gggcagagag tggactacg ggaacgtgac 900  
 gcggcaccag gcgcagtgcg ccctcaacga gcggggcggtg gagggcgact tcctcattag 960  
 ggacagcgag tcctcgccca gcgacttctc cgtgtccctt aaagcgtcag ggaagaacaa 1020  
 acacttcaag gtgcagctcg tggacaatgt ctactgcatt gggcagcggc gcttccacac 1080  
 catggacgag ctggtggaac actacaaaaa ggcgccccatc ttcaccagcg agcacgggga 1140  
 gaagctctac ctcgtcaggg ccctgcagtg acggcgcccc ggcccacac tcgcctcccg 1200  
 ggcccacggt tggagctgcc cgcccggcct tgtggcagag gctcctcccg cggggacggc 1260  
 cccgacggct tctctg 1276

<210> 22

<211> 1577

<212> DNA

<213> NM\_006214.2| Homo sapiens phytanoyl-CoA hydroxylase (Refsum disease)  
 (PHYH), mRNA

<400> 22  
 gccgcgtgcg gtaaatgggg cagaggccgg gaggggtggg ggttccccgc gccgcagcca 60  
 tggagcagct tcgcgccgcc gcccgctctgc agattgttct gggccacctc ggccgccccct 120  
 cggccggggc tgtcgtagct catccactt cagggactat ttcctctgcc agtttccatc 180  
 ctcaacaatt ccagtatact ctggataata atgttctaac cctggaacag agaaaatttt 240  
 atgaagaaaa tgggtttcta gtaatcaaaa atcttgtacc tgatgccgat attcaacgct 300  
 ttcggaatga gtttgaaaaa atctgcagaa aggaggtgaa accattagga ttaacagtaa 360  
 tgagagatgt gaccatttcg aaatccgaat atgctccaag tgagaagatg atcacgaagg 420  
 tccaggattt ccaggaagat aaggagctct tcagatactg cactctcccc gagattctga 480  
 aatatgtgga gtgcttact ggacctaata ttatggccat gcacacaatg ttgataaaca 540  
 aacctccaga ttctggcaag aagacgtccc gtcacccct gcaccaggac ctgcactatt 600  
 tccccctcag gcccgagcgt ctcctcgtt gcgcctggac ggcgatggag cacatcagcc 660  
 ggaacaacgg ctgtctgggt gtgtctccag gcacacacaa gggctccctg aagccccacg 720  
 attaccccaa gtgggagggg ggagttaaca aaatgttcca cgggatccag gactacgagg 780  
 aaaacaaggc ccgggtgcac ctggtgatgg agaagggcga cactgttttc ttccatcctt 840

tgctcatcca	cggatctggt	cagaataaaa	cccagggatt	ccggaaggca	atttcctgcc	900
atttcgccag	tgccgattgc	cactacattg	acgtgaaggg	caccagtcaa	gaaaacatcg	960
agaaggaagt	tgtaggaata	gcacataaat	tctttggagc	tgaaaatagc	gtgaacttga	1020
aggatatttg	gatgtttcga	gctcgacttg	tgaaaggaga	aagaaccaat	ctttgaaata	1080
gccatctgct	ataactcttt	caacagaaaa	ccaaaaccaa	acgaaatgtc	taaggaaaat	1140
gttttcttaa	tgagatgatg	taaccttttc	tatcacttgt	taaaagcaga	aaacatgtat	1200
caggctactta	attgcataga	gttagttttg	cagcacaatg	gtgttgcttt	aatggaaaaa	1260
aaaaacagta	aaagtgaaat	attactgttt	taaggaaaac	taatttaggg	tggcagccaa	1320
taaagggtggt	tgggtgtctaa	tttaagtgtt	aaatcaatth	ctttcattca	gttagctctt	1380
tacccaagaa	gaagtgaatg	atttggagct	taggggtatgt	tttgtatccc	ctttctgata	1440
aaccatttcc	ctaccaatth	tatgtcataa	gagattttth	tcccccaaath	ctagaacaath	1500
gtataatata	ttcacatcta	gtcaagggca	taggaacggt	gtcatggagt	ccaaataaag	1560
tgatatttcc	tgctcgg					1577

<210> 23

<211> 3060

<212> DNA

<213> NM\_004739.2| Homo sapiens metastasis-associated gene family, member 2 (MTA2), mRNA

<400> 23						
tccggaagga	ggcgaaccct	gaggcggggc	cggcaagcct	tccctgcggc	cggcagagcc	60
caacgactag	tgggactccg	cgggggcggg	ggtagctgga	gcctggctct	ggcctggcag	120
gagccgagct	tgttccgga	gaagccgagc	ggacgggggc	cagcctcagc	gtcccgggag	180
tgaggcgata	gctgcggcgg	cgacagcgcg	ggccgggatg	aaccgcgacg	gctgaggcag	240
cggaggtgcc	ggctgcgcgg	gccccagtga	gactccctcg	aagcggcagc	ccaccgttcg	300
gggctttgcc	tcgagccgag	ccctgcccc	gcgagcctcc	cggacccctt	tgtgcggccg	360
gaggcggcgg	cgggaacggc	catggcggcc	aacatgtacc	gggtgggaga	ttacgtctat	420
tttgagaact	cttccagcaa	tccttacctg	gttagacgga	ttgaggagct	caacaagact	480
gcaaattgga	atgtggaggc	aaaggttgtc	tgtcttttcc	ggcgcaggga	catttctagt	540
agcctcaaca	gcctggctga	tagtaatgcc	agggagtth	aagaggaatc	aaagcagcca	600
ggggtgtctg	agcagcagcg	ccatcaactg	aagcaccggg	aactttttct	ttctcggcaa	660
tttgaatcat	taccagccac	ccacatacgg	gggaaatgca	gtgtgaccct	cttgaatgag	720
acagatatct	tgagccagta	cctggaaaag	gaggactgct	ttttttactc	actgggtgtt	780
gaccccgctg	agaagacact	tctcgtgat	cagggcgaga	ttagagttgg	ttgcaaatac	840



caagctgaga tcccagatcg cctagtagag ggagaatctg ataatcggaa ccagcagaag	900
atggagatga aggtctggga cccagacaac cctctcacag accggcagat cgaccagttt	960
cttgtggtgg cccgagctgt gggaaccttt gcaagagccc tagattgtag cagctccatt	1020
cggcagccaa gcttgccacat gagtgcagct gctgcctccc gagatatcac tctgtttcac	1080
gccatggata ccttgcaaaag gaacggctac gacctggcta aggccatgtc gaccctggta	1140
ccccagggag gcccgggtgct gtgtcgggat gagatggagg aatggtcagc ctcagaggcc	1200
atgctatttg aggaggccct agagaagtat gggaaggact tcaatgatat tcgccaggat	1260
tttctacctt ggaagtcact tgccagcata gtccagtttt attacatgtg gaaaaccaca	1320
gaccggtata ttcagcagaa aagggtgaaa gctgctgaag cagacagcaa actgaaacag	1380
gtctacattc ccacctacac taagccaaac cctaaccaga tcatttctgt gggttcaaaa	1440
cctggcatga atggggctgg atttcagaag ggcctgactt gtgagagttg ccacaccaca	1500
cagtctgctc agtggatatgc ctggggccca cctaacatgc agtgccgcct ctgtgcttcc	1560
tgttgatct actggaagaa gtatggggga ctgaagaccc caactcagct tgagggggcc	1620
actcggggca ccacggagcc aactcaagg ggtcatttat ccagacctga agctcaaagt	1680
ctctctcctt acacaaccag cgccaacagg gccaaagctac tggctaagaa cagacaaact	1740
ttcctgcttc agaccacaaa gctgacctgt cttgccagac gcatgtgcag ggacctatta	1800
cagccaagga gggccgccc acggccttat gtcctatca atgccaatgc catcaaagca	1860
gagtgtcca ttcgacttcc taaggccgcc aagactccat tgaagattca ccctctggtg	1920
cggctgcccc tggcaactat cgtcaaagat ctggtggccc aggcaccctt gaaacaaaaa	1980
acacctcggg gtaccaagac accgatcaac agaaaccagc tgtcccagaa ccggggactg	2040
gggggcatta tggtgaaacg ggcctatgag actatggcag gggcaggggt tcctttctct	2100
gccaatggaa ggcctctggc ttcagggatt cgttcaagct cacagccagc agccaagcgt	2160
cagaaactaa acccagctga tgcccccaat cctgtggtgt ttgtggccac aaaggatacc	2220
agggccctac ggaaggctct gacctatctg gaaatgcggc gagctgctcg ccgacctaac	2280
ttgcccctga aggtgaagcc aacgctgatt gcagtgcggc cccctgtccc tctacctgca	2340
ccctcacatc ctgccagcac caatgagcct attgtcctgg aggactgagc acctgtgggg	2400
aaggaggtg ggctgagagg tagagggtg atgcccaggg cacccaaacc tcccttccct	2460
ttcgtgtcga agggagtgag gagtgaatta aggaagagag caagtgagtg tgtgtccctg	2520
gaggggttg gcgcccctct gtgttaccac ctcgagactt gtctcatgcc tccatgcttg	2580
ccgatggagg acagactgca ggaacttggc ccatgtggga acctagcctg ttttgggggg	2640
taggacctac agatgtcttg gacagttttg gggggagggg tttttaattt tttaaaagtt	2700
ttgcctccct ttgtgaaagg ggatggggag gggaagagta aacagataac aggtgggtgg	2760
acctggttg gggagggggg cgtgcactgc catgtctttt tttttttttt tttttttttt	2820

tttcctaatt	gggggtttct	ctttctgtcc	ggtgtccgga	ctttcctaatt	tggagtttga	2880
ggcccctaag	ctggcatcaa	ccccaggcca	cgctcgctct	ttccttccct	cccctcccc	2940
tctgcctttt	gtacgccagt	tctcagaaat	aaagatcttt	tgtccgtttt	tttaacctcg	3000
gattctgtaa	ttggttctta	tagtaacaaa	taaaaagctg	ttttcttcag	cttctcctgg	3060

<210> 24

<211> 2407

<212> DNA

<213> NM\_001091.1| Homo sapiens amiloride binding protein 1 (amine oxidase (copper-containing)) (ABP1), mRNA

<400> 24		
gatacagctta	agggcaaagg	ctggaagcag agcgaactgg gagcagagca cacagagccg 60
tggagcgaga	gatgccggcc	ctgggctggg ccgtggctgc catcctgatg ctgcagacgg 120
ccatggcgga	gccctccccg	gggactctgc ccaggaaggc aggggtgttt tcagacctaa 180
gcaaccaaga	gctgaaggca	gtgcacagct tcctctgggtc caagaaggag ctgaggctgc 240
agccctccag	taccaccacc	atggccaaga acaccgtgtt tctcatcgag atgctgctgc 300
ccaagaagta	ccatgtgctg	aggtttcttg ataaagggtga aaggcatcct gtgcgggaag 360
cccgtgccgt	catcttcttt	ggtgaccagg agcatcccaa tgtcaccgag tttgctgtgg 420
ggccccctgcc	agggccccctgc	tacatgcgag cactgtcccc caggcctggg taccagtcct 480
cctgggcatc	gaggcccatc	tccacagcag agtatgccct cctctaccac accctgcagg 540
aagccaccaa	gccccctgcat	cagttcttcc tcaataccac aggtttctca ttccaagact 600
gccatgacag	atgcctggcc	ttcaccgatg tggccccccg ggggtgtggct tctggccagc 660
gccgcagttg	gcttatcata	cagcgctatg tagaaggcta ctttctgcac cccactgggc 720
tggagctcct	cgtaggatcat	gggagcacag atgctgggca ctgggccgtg gagcagggtgt 780
ggtacaacgg	gaagttctat	gggagcccag aggaactggc tcggaagtat gcagatggag 840
aggtggacgt	ggtggtcctg	gaggacccgc tgcctggggg caaggggcat gacagcacag 900
aggagccgcc	cctcttctcc	tcccacaagc cccgcgggga cttccccagc cccatccatg 960
tgagcggccc	ccgcttggtc	cagccccacg gccctcgctt caggctggag ggcaacgctg 1020
tgctctacgg	cggctggagc	tttgccctcc ggctgcgctc ctccctccggg ctgcagggtcc 1080
tgaacgtgca	cttcggcgga	gagcgcattg cctatgaggt cagcgtgcaa gaggcagtggt 1140
cgctgtatgg	aggacacaca	cctgcaggca tgcagaccaa gtacctcgat gtcggctggg 1200
gcctgggcag	cgtcactcat	gagttagccc ccggcatcga ctgcccggag accgccacct 1260
tcctggacac	tttccactac	tatgatgccg atgaccgggt ccattatccc cgagccctct 1320
gcctctttga	aatgcccaca	ggggtgcccc ttcggcggca ctttaattcc aactttaaag 1380

gtggcttcaa cttctatgcg gggctgaagg gccaggtgct ggtgctgcgg acaacttcaa	1440
ctgtctacaa ttatgattac atttgggact ttatcttcta cccaacggg gtgatggagg	1500
ccaagatgca tgccactggc tacgtccacg ccaccttcta ccccccgag gggctgcgcc	1560
acggcactcg cctgcacacc cacctgattg gcaacataca cactcacttg gtgcactacc	1620
gcgtagacct ggatgtggca ggcaccaaga acagcttcca gacactgcag atgaagctag	1680
aaaacatcac caacccttg agcccaagac accgctgggt ccagccaact ctggagcaga	1740
cgcagtactc ctgggagcgc caggcggcct tccgcttcaa aaggaagctg cccaagtacc	1800
tgctctttac cagccccag gagaaccctt ggggccacaa gcgcagctac cgcctgcaga	1860
tccactccat ggccgaccag gtgctgcccc caggctggca ggaggagcag gccatcacct	1920
gggcaaggta ccccttgga gtgaccaagt accgggagtc ggagctgtgc agcagcagca	1980
tctaccacca gaacgacccc tggcaccgc ccgtggtctt tgagcagttt cttcacaaca	2040
acgagaacat tgaaaatgag gacctggtgg cctgggtgac ggtgggcttc ctgcacatcc	2100
cccactcaga ggacattccc aacacagcca cacctgggaa ctccgtgggc ttcctgctcc	2160
ggccattcaa cttcttccca gaggaccctt ccctggcatc cagagacact gtgatcgtgt	2220
ggcctcggga caacggcccc aactacgtcc agcgtggat ccctgaggac agggactgct	2280
cgatgcctcc cccttttagc tacaatggga cctatagacc tgtgtgacca gccccagtt	2340
cctccccag ttcctcccag gaagcccagg agcctcactg gggcagacaa taaactctca	2400
gagcctc	2407

<210> 25

<211> 1094

<212> DNA

<213> NM\_000712.3| Homo sapiens biliverdin reductase A (BLVRA), mRNA

<400> 25	
ccgccccggt ccgcaaagcc ggtggcgccc ggaggctgca cggagagcgg tgcccgcgtc	60
agtgaccgaa ggaagagacc aagatgaatg cagagcccga gaggaagttt ggcgtggtgg	120
tggttggtgt tggccgagcc ggctccgtgc ggatgagggga cttgcggaat ccacaccctt	180
cctcagcgtt cctgaacctg attggcttcg tgtcgagaag ggagctcggg agcattgatg	240
gagtcacgca gatttctttg gaggatgctc tttccagcca agaggtggag gtcgcctata	300
tctgcagtga gagctccagc catgaggact acatcaggca gttccttaat gctggcaagc	360
acgtccttgt ggaatacccc atgacactgt cattggcggc cgctcaggaa ctgtgggagc	420
tggctgagca gaaaggaaaa gtcttgacag aggagcatgt tgaactcttg atggaggaat	480
tcgctttcct gaaaaaagaa gtggtgggga aagacctgct gaaagggctg ctctcttca	540
cagctggccc gttggaagaa gagcggtttg gcttcctgc attcagcggc atctctcgcc	600

tgacctggct ggtctccctc tttggggagc tttctcttgt gtctgccact ttggaagagc	660
gaaaggaaga tcagtatatg aaaatgacag tgtgtctgga gacagagaag aaaagtccac	720
tgtcatggat tgaagaaaaa ggacctgggc taaaacgaaa cagatattta agcttccatt	780
tcaagtctgg gtccttggag aatgtgccaa atgtaggagt gaataagaac atatttctga	840
aagatcaaaa tatatttgtc cagaaactct tgggccagtt ctctgagaag gaactggctg	900
ctgaaaagaa acgcatcctg cactgcctgg ggcttgaga agaatccag aaatattgct	960
gttcaaggaa gtaagaggag gaggtgatgt agcacttcca agatggcacc agcatttggg	1020
tcttctcaag agttgacctat tatctctatt cttaaaatta aacatgttgg ggaaacaaga	1080
aaaaaaaaaa aaaa	1094

<210> 26

<211> 5546

<212> DNA

<213> NM\_000933.2| Homo sapiens phospholipase C, beta 4 (PLCB4), transcript variant 1, mRNA

<400> 26	
aggaaaagac aatttctctc tgattcagaa tcctgaaaat gtgatctccc ttaaaaagag	60
gacagtgctg ctgtgagttt gacgaagtgg acatcacctg cagtcagtcc agagctgccc	120
agtcttgaat ataatcatgg ccaaacctta tgaatttaac tggcagaagg aagttccctc	180
ctttttgcaa gaaggaacag tttttgacag atacgaggag gaatcctttg tgtttgaacc	240
caactgcctc ttcaaagtgg atgagtttgg cttctttctg acatggagaa gtgaaggcaa	300
ggaaggacag gtgctagaat gctccctcat caacagtatt cggtcgggag ccataccaaa	360
ggatcccaaa atcttggctg ctcttgaagc tgttggaaaa tcagaaaatg atctggaagg	420
gcggatagtt tgtgtctgca gtggcacaga tctagtgaac attagtttta cctacatggt	480
ggctgaaaat ccagaagtaa ctaagcaatg ggtagaaggc ctgagatcaa tcatacacia	540
cttcagggcc aacaacgtca gtccaatgac atgcctcaag aaacactgga tgaaattggc	600
atttatgacc aacacaaatg gtaaaattcc agttaggagt attactagaa catttgcatc	660
gggaaaaaca gaaaagggtga tctttcaagc actcaaggag ttaggtcttc ccagtggaaa	720
gaatgatgaa attgagccca cagcattttc ttatgaaaag ttctatgaac tgacacaaaa	780
gatttgtcct cggacagata tagaagatct tttcaaaaaa atcaatggag acaaaaactga	840
ttattttaacg gtagaccaat tagtgagctt tctaaatgaa catcaacgag atcctcgatt	900
gaatgaaatt ttatttccat tttatgatgc caaaagggca atgcagatca ttgagatgta	960
tgaacctgat gaagatttga agaaaaaagg ccttatatca agtgatgggt tttgcagata	1020
tctgatgtca gatgaaaacg ccccagtcct cctagatcgt ttagaacttt accaagaaat	1080

ggaccatcct ctggctcact acttcatcag ttcttcccat aacacttata tcaactggcag	1140
acagttcggc gggaagtctt cggtagaaat gtacagacag gttctcctgg ctggttgacg	1200
atgtgttgaa cttgactgct gggatggaaa aggtgaagac caagaaccaa taataactca	1260
tggaaaagca atgtgtacag atatcctttt taaggatgta attcaagcca tcaaggaaac	1320
tgcatttgct acatcagaat atcctgtaat tctctccttt gaaaatcact gcagcaaata	1380
tcaacagtac aagatgtcca aatattgcga agatctatctt ggggatctcc tgttgaaaca	1440
agcacttgaa tcacatccac ttgaaccagg cagggtcttg ccatcccca atgacctcaa	1500
aagaaaaata ctcataaaaa acaagcggct gaaacctgaa gttgaaaaaa aacagctgga	1560
agctttgaga agcatgatgg aagctggaga atctgcctcc ccagcaaaca tcttagagga	1620
cgataatgaa gaggagatcg aaagtgtga ccaagaggag gaagctcacc ccgaattcaa	1680
atgtggaaat gaactttctg ctgatgactt gggtcacaag gaagctgttg caaatagcgt	1740
caagaagggc ctggctcactg tagaagatga gcaggcgtgg atggcatctt ataaatatgt	1800
aggtgctacc actaatatcc atccatattt gtccacaatg atcaactacg cccagcctgt	1860
aaagtttcaa ggtttccatg tggcagaaga acgcaatatt cattataaca tgtcttcttt	1920
taatgaatca gtcggctctg gctacttgaa gacacatgca attgaatttg tcaattataa	1980
caaacggcaa atgagtcgca tttaccccaa gggaggccga gtcgattcca gtaattacat	2040
gcctcagatt ttctggaacg ctggctgcca gatgggttca ctgaactatc aaaccccgaa	2100
tttagcgatg caattgaatc agggaaaatt tgagtataat ggatcgtgcg ggtaccttct	2160
caaaccagat ttcattgaggc ggcctgatcg aacatttgac cccttctctg aaactcctgt	2220
tgatgggtgtt attgcagcca cttgctcagt gcagggtata tcagggtcaat tcttatcaga	2280
taagaaaatt ggcacctacg tagagggtgga tatgtatggg ttgcccactg acaccatacg	2340
taaggaattc cgaactcgca tgggttatgaa taatggactc aatccagttt acaatgaaga	2400
gtcatttgta tttcggaagg tgatcctgcc ggacctggct gtcttgagaa tagctgtgta	2460
tgatgataac aacaagctga ttggccagag gatcctcccg cttgatggcc tccaagccgg	2520
atatcgacac atttcccttc gaaatgaggg aaataaacca ttatcactac caacaatttt	2580
ctgcaatatt gttcttaaaa catatgtgcc tgatggattt ggagatatcg tggatgcttt	2640
atcagatcca aagaaatttc tctcaattac agaaaagaga gcagaccaa tgagagctat	2700
gggcattgaa actagtgaca tagccgacgt gccagtgac acttccaaaa atgacaagaa	2760
aggaaaggcc aacaccgcca aagcaaatgt gaccctcag agtagctctg agctcagacc	2820
aaccaccacg gctgccctgg cctctgggtg ggaagccaag aaaggatttg aacttatccc	2880
tcaagtaagg atagaagact taaagcagat gaaggcttac ttgaagcatt taaagaaaca	2940
gcagaaggag ctaaattctt taaagaagaa acatgcaaag gaacacagta ccatgcagaa	3000
gttacactgc acgcaagttg acaaaattgt ggcacagtat gacaaagaga agtcgactca	3060

tgagaaaatc ctagagaagg caatgaagaa gaagggggga agtaattgtc tcgaaatgaa	3120
aaaagaaaca gaaatcaaaa ttcagacgct gacatcagat cacaaatcta aggtcaaaga	3180
gattgtagca cagcacacaa aggaatggtc agaaatgata aatacccaca gtgctgagga	3240
gcaagaaatc cgagacctgc acctcagcca gcagtgtgag ctgctgaaaa agctactcat	3300
caatgcccac gagcagcaaa cccagcagct gaaactgtcc catgacaggg aaagcaagga	3360
aatgcgagca caccaggcta agatttctat ggaaaatagc aaagccatca gccaagataa	3420
atctatcaag aataaagcag aacgggaaag gcgagtcagg gagttaaaca gcagcaacac	3480
taaaaagttt ctggaagaaa gaaagagact tgccatgaag cagtccaaag aaatggatca	3540
gttgaaaaaa gtccagcttg aacatctaga attcctagag aaacagaatg agcagctttt	3600
gaaatcctgt catgcagtgt cccaaacgca aggcgaagga gatgcagcag atggtgaaat	3660
tggaagccga gatggaccgc agaccagcaa cagtagtatg aaactccaaa atgcaaactg	3720
aagcagcaaa cccacaaagc atcaaaaagac tcaactcaca acttctgaac acaaactcca	3780
tggatgaaag ctgtttatit tgtttccttt atgtgtaaac aagatgatat ctgaaaccag	3840
agagacttgg aatgtctgac tgacttctat ttaacagctt gagtattgca tttccttggc	3900
caaacaaaaa tagctacaaa tccacaaaaa tttactattc cagtaaggca gagtccaacc	3960
attgataata caacttaaac atgtttgcta taaaatacca tcacaagtaa atgagcttgg	4020
tgtgaacaac tctcctttgt gatgccttag gacatgtttg aactgcagca aaaaacaaaa	4080
acaaaaaaca gtgcattagc aatttcatag caagtgcag cactaggaaa agaaaactct	4140
gtctacaagt ttattagcag aagtgggtgt ctgctagaca aataattttg caaaattttt	4200
ctacatctaa gttacctcat cagtaagtgc catgtctcta ccatgccatc agaggcta	4260
ttcctgtaaa agttgtggaa attgttagaa caatagaaaa atagagcagt gtatgtgtgc	4320
caaaactcat cattactcaa aggagaactg tgttaggcac atttaagaaa gtttacctt	4380
gacattgctt tataggaatt gtttctgcag attccggata ttataattca caccataaag	4440
attgtgaagt ggttattggc aaacgtttgt aaatgtgacc atgtataaag tatttatact	4500
ctttaattca cactgttaga gagcaaaatc atctaagtat tgccacatga caagattagt	4560
aaacaggaat actagaacta tgtttgcatg atacacaagc accaataaag actaatccat	4620
acacagttaa cctaattgcca aataaatact ggttaaataa atgtatggca cagaatataa	4680
tttgactatc aagactttta gcataatgaa aaacctctc tctatatata tatgtgtata	4740
tgaattatgt gggcattctt gatacttcaa gttctagttt gaaaaaaaata cataactaat	4800
ttaatattac acaaaaatat ttatgcagat tttcagaatt tcatatcagg aaatgacctt	4860
tttatgtctg ttaaatatca aaacaatttg ctacagtgtt aatctgcatg gtctttaagc	4920
ctgctgtagt tgagttgcag acagtgcag aaaaagtatt ccgctgggaa ttgagccatg	4980
ccaccaaaag caagaggagc gcatggaaac ccggtagtct agaactaatc agattactga	5040
ttttagggca cagcaccaga tgaattgttg tatatgcttg taaaaattga ttctgtgtgt	5100

tcctctgaac aaagcggaga aaatgatgat accatcaata ttgaaattaa acttccaact	5160
tctctaataa aaaattaaaa caccgataac actcgtcaag agtatttgct cccaagacac	5220
attctagcaa atgttttgcc tttttcatat acatgatatc atcgttattt tcaaaggggg	5280
cttattaata ccttcagcat gtttttcacc caaatgatgc aaaacatgca gattctagtt	5340
gacttcagtt gtaatagact tgtttttctc ctatttatga tttgaagtgg attctgtaaa	5400
atatctcttg ttcttagttt ccttatctgt aaaacagtgg agtttagacta catatctttt	5460
ggcactaaca tctcatgaaa aattatgggt aataaaatat caccacattt ggattgccaa	5520
ttttcaaaaa aaaaaaaaaa aaaaaa	5546

<210> 27

<211> 2545

<212> DNA

<213> NM\_002416.1| Homo sapiens chemokine (C-X-C motif) ligand 9 (CXCL9), mRNA

<400> 27	
atccaataca ggagtgactt ggaactccat tctatcacta tgaagaaaag tggtgttctt	60
ttctcttgg gcatcatctt gctggttctg attggagtgc aaggaacccc agtagtgaga	120
aagggtcgct gttcctgcat cagcaccaac caagggacta tccacctaca atccttgaaa	180
gaccttaaac aatttgcccc aagcccttcc tgcgagaaaa ttgaaatcat tgctacactg	240
aagaatggag ttcaaacatg tctaaacca gattcagcag atgtgaagga actgattaaa	300
aagtgggaga aacaggtcag ccaaaagaaa aagcaaaaga atgggaaaaa acatcaaaaa	360
aagaaagttc tgaaagttcg aaaatctcaa cgttctcgtc aaaagaagac tacataagag	420
accacttcac caataagtat tctgtgttaa aaatgttcta ttttaattat accgctatca	480
ttccaaagga ggatggcata taatacaaag gcttattaat ttgactagaa aatttataaac	540
attactctga aattgtaact aaagttagaa agttgatttt aagaatccaa acgttaagaa	600
ttgttaaagg ctatgattgt ctttggttctt ctaccaccca ccagttgaat ttcacatgc	660
ttaaggccat gattttagca atacctatgt ctacacagat gttcacccaa ccacatccca	720
ctcacaacag ctgcctggaa gagcagccct aggcttcac gtactgcagc ctccagagag	780
tatctgaggc acatgtcagc aagtcctaag cctgttagca tgctggtgag ccaagcagtt	840
tgaaattgag ctggacctca ccaagctgct gtggccatca acctctgtat ttgaatcagc	900
ctacaggcct cacacacaat gtgtctgaga gattcatgct gattgttatt gggatcacc	960
actggagatc accagtgtgt ggctttcaga gcctcctttc tggctttgga agccatgtga	1020
ttccatcttg cccgctcagg ctgaccactt tatttctttt tgttccccct tgcttcattc	1080
aagtcagctc ttctccatcc taccacaatg cagtgccttt cttctctcca gtgcacctgt	1140

catatgctct gatttatctg agtcaactcc tttctcatct tgtccccaac accccacaga	1200
agtgccttct tctcccaatt catcctcact cagtcagct tagttcaagt cctgcctctt	1260
aaataaacct ttttggacac acaaattatc ttaaaactcc tgtttcactt ggttcagtag	1320
cacatgggtg aacactcaat ggttaactaa ttcttgggtg tttatcctat ctctccaacc	1380
agattgtcag ctcttgagg gcaagagcca cagtatatct ccctgtttct tccacagtgc	1440
ctaataatac tgtggaacta ggttttaata attttttaat tgatgttggt atgggcagga	1500
tggcaaccag accattgtct cagagcaggt gctggctctt tcctggctac tccatgttgg	1560
ctagcctctg gtaacctctt acttattatc ttcaggacac tcactacagg gaccagggat	1620
gatgcaacat ccttgtcttt ttatgacagg atgtttgctc agcttctcca acaataagaa	1680
gcacgtggta aaacacttgc ggatattctg gactgttttt aaaaaatata cagtttaccg	1740
aaaatcatat aatcttacia tgaaaaggac tttatagatc agccagtgc caaccttttc	1800
ccaaccatac aaaaattcct tttcccgaag gaaaagggtt ttctcaataa gcctcagctt	1860
tctaagatct aacaagatag ccaccgagat ccttatcgaa actcatttta ggcaaatatg	1920
agttttattg tccgtttact tgtttcagag ttgtattgt gattatcaat taccacacca	1980
tctcccatga agaaaggga cgggtgaagta ctaagcgcta gaggaagcag ccaagtcggt	2040
tagtggaagc atgattgggtg ccagtttagc ctctgcagga tgtggaaacc tccttccagg	2100
ggaggttcag tgaattgtgt aggagaggtt gtctgtggcc agaattttaa cctatactca	2160
ctttcccaa ttgaatcact gctcacactg ctgatgattt agagtgtgt ccggtggaga	2220
tcccaccga acgtcttatt taatcatgaa actccctagt tccttcatgt aacttccctg	2280
aaaaatctaa gtgtttcata aatttgagag tctgtgacct acttaccttg catctcacag	2340
gtagacagta tataactaac aaccaaagac tacatattgt cactgacaca cacgttataa	2400
tcatttatca tatatataca tacatgcata cactctcaaa gcaaataatt tttcacttca	2460
aaacagtatt gacttgtata ccttgaatt tgaaatattt tctttgttaa aatagaatgg	2520
tatcaataaa tagaccatta atcag	2545

<210> 28

<211> 1144

<212> DNA

<213> NM\_005859.2| Homo sapiens purine-rich element binding protein A (PURA), mRNA

<400> 28

cgactgaggc ggcgggcgga gcggcaggcg gcggcggcgc ggcagcggag cgcagcatca	60
tggcggaccg agacagcggc agcgagcagg gtggtgcggc gctgggttcg ggcggctccc	120
tggggcacc cggctcgggc tcaggctccg gcgggggcgg tggtggcggc gggggcggcg	180



gcggcagtgg	cggcggcggc	ggcggggccc	caggggggct	gcagcacgag	acgcaggagc	240
tggcctccaa	gcgggtggac	atccagaaca	agcgcttcta	cctggacgtg	aagcagaacg	300
ccaagggccg	cttcctgaag	atcgccgagg	tgggcgcggg	cggaacaag	agccgcctta	360
ctctctccat	gtcagtggcc	gtggagtccc	gcgactacct	gggcgacttc	atcgagcact	420
acgcgcagct	gggccccagc	cagccgccgg	acctggccca	ggcgcaggac	gagccgcgcc	480
gggcgctcaa	aagcgagtcc	ctggtgcgcg	agaaccgcaa	gtactacatg	gatctcaagg	540
agaaccagcg	cgcccgcttc	ctgcgcatcc	gccagacggt	caaccggggg	cctggcctgg	600
gctccacgca	gggccagacc	attgcgctgc	ccgcgcaggg	gctcatcgag	ttccgtgacg	660
ctctggccaa	gctcatcgac	gactacggag	tggaggagga	gccggccgag	ctgcccgagg	720
gcacctcctt	gactgtggac	aacaagcgct	tcttcttcga	tgtgggctcc	aacaagtacg	780
gcgtgtttat	gcgagtgagc	gaggtgaagc	ccacctatcg	caactccatc	accgtgccct	840
acaaggtgtg	ggccaagtcc	ggacacacct	tctgcaagta	ctcggaggag	atgaagaaga	900
ttcaagagaa	gcagagggag	aagcgggctg	cctgtgagca	gcttcaccag	cagcaacagc	960
agcagcagga	ggagaccgcc	gctgccaccc	tgctactgca	gggtgaggaa	gaaggggaag	1020
aagattgatc	aaactgaatg	aaacccccac	acacacacac	atgcatacac	acacacacac	1080
agccacacac	acagaaaata	tactgtaaag	aaagagagaa	aataaaaagt	taaaaagtta	1140
aaaa						1144

<210> 29

<211> 1575

<212> DNA

<213> NM\_014298.3| Homo sapiens quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating)) (QPRT), mRNA

<400> 29	
tcccaccccc	agcctggggc
ctctgggagc	cttggctctg
agcagccaac	acaccagccc
60	
agacagctgc	aagtcaccat
ggacgctgaa	ggcctggcgc
tgctgctgcc	gcccgtcacc
120	
ctggcagccc	tggaggagag
ctggctccga	gaggactgcc
cagggctcaa	ctacgcagcc
180	
ttggtcagcg	gggcaggccc
ctcgaggcgc	gcgctgtggg
ccaaatcccc	tgggggtactg
240	
gcagggcagc	ctttcttcga
tgccatattt	acccaactca
actgccaaagt	ctcctggttc
300	
ctccccgagg	gatcgaagct
ggtgccgggtg	gccagagtgg
ccgagggtccg	gggccctgcc
360	
cactgcctgc	tgctggggga
acgggtggcc	ctcaacacgc
tgccccgctg	cagtggcatt
420	
gccagtgctg	ccgccgctgc
agtggaggcc	gccagggggg
ccggctggac	tgggcacgtg
480	
gcaggcacga	ggaagaccac
gccaggcttc	cggctggtgg
agaagtatgg	gctcctgggtg
540	
ggcggggccg	cctcgaccgc
ctacgacctg	ggagggctgg
tgatggtgaa	ggataaccat
600	

gtggtggccg cccgtggcgt ggagaaggcg gtgcgggagg ccagacaggc ggctgacttc	660
gctctgaagg tgggaagtgga atgcagcagc ctgcaggagg ccgtgcaggc agctgaggct	720
ggtgccgacc ttgtcctgct ggacaacttc aagccagagg agctgcaccc cacggccacc	780
gtgctgaagg cccagttccc gagtgtggct gtggaagcca gtgggggcat caccctggac	840
aacctcccc agttctgcgg gccgcacata gacgtcatct ccatggggat gctgaccag	900
gcggccccag cccttgattt ctccctcaag ctgtttgcc aagagggtggc tccagtggcc	960
aaaatccact agtcctaaac cggaagagga tgacaccggc catgggttaa cgtggctcct	1020
caggaccctc tgggtcacac atctttagggt tcagtggcca atggggcaca tttggcacta	1080
gcttgagccc aactctggct ctgccacctg ctgctcctgt gacctgtcag ggctgacttc	1140
acctctgtct atctcagttt cctaattctgt aaaatgggtc taataaagga tcaaccacat	1200
ggggttctgc ggtgataatg agcacatagt gaggggtcag caaatgtcag aagttacctg	1260
ggacagccgg gcacgatggc tcacacctgt aatcccagca ctttgggagg ctgaggcggg	1320
aagatcactt gagttcagga gtttgagacc agcctggcca acatggtgaa accccatctc	1380
tacaaaaaat agaagaatta gctgggtgtg gtggcacgcg cctgtaatcc cagctactta	1440
ggaggctgag gcaggagaat cgcttgaacc caggaagtgg aggttgcagt gagctgatgg	1500
tgccactgca ctccagcctg ggtgatagag cgagactctg tctccaaaga agaaaaaaaa	1560
aaaaaaaaaa aaaaa	1575

<210> 30

<211> 768

<212> DNA

<213> NM\_004585.2| Homo sapiens retinoic acid receptor responder (tazarotene induced) 3 (RARRES3), mRNA

<400> 30	
ccttcagcat aaaagctgat ccacaaacaa gaggagcacc agacctcctc ttggcttcga	60
gatggcttcg ccacaccaag agcccaaacc tggagacctg attgagattt tccgccttgg	120
ctatgagcac tgggccctgt atataggaga tggctacgtg atccatctgg ctctccaag	180
tgagtacccc ggggctggct cctccagtgt cttctcagtc ctgagcaaca gtgcagaggt	240
gaaacggggg cgcctggaag atgtggtggg aggctgttgc tatcgggtca acaacagctt	300
ggaccatgag taccaaccac ggcccgtgga ggtgatcatc agttctgcga aggagatggt	360
tggtcagaag atgaagtaca gtattgtgag caggaactgt gagcactttg tcgccagct	420
gagatatggc aagtcccgtg gtaaacaggt ggaaaaggcc aaggttgaag tcggtgtggc	480
cacggcgctt ggaatcctgg ttgttgctgg atgctctttt gcgattagga gataccaaaa	540
aaaagcaaca gcctgaagca gccacaaaat cctgtgttag aagcagctgt gggggtccca	600

gtggagatga gcctccccc tgcctccagc agcctgaccc tcgtgccctg tctcaggcgt 660  
tctctagatc ctttcctctg tttccctctc tcgctggcaa aagtatgac taattgaaac 720  
aagactgaag gatcaataaa cagccatctg ccccttcaaa aaaaaaaaa 768

<210> 31

<211> 696

<212> DNA

<213> NM\_002984.1| Homo sapiens chemokine (C-C motif) ligand 4 (CCL4), mRNA

<400> 31  
ttccccccc ccccccccc ccccgcccga gcacaggaca cagctgggtt ctgaagcttc 60  
tgagttctgc agcctcacct ctgagaaaac ctcttttcca ccaataccat gaagctctgc 120  
gtgactgtcc tgtctctcct catgctagta gctgccttct gctctccagc gctctcagca 180  
ccaatgggct cagaccctcc caccgcctgc tgcttttctt acaccgcgag gaagcttcct 240  
cgcaactttg tggtagatta ctatgagacc agcagcctct gctcccagcc agctgtggta 300  
ttccaaacca aaagaagcaa gcaagtctgt gctgatccca gtgaatcctg ggtccaggag 360  
tacgtgtatg acctggaact gaactgagct gctcagagac aggaagtctt cagggaaggt 420  
cacctgagcc cggatgcttc tccatgagac acatctcctc catactcagg actcctctcc 480  
gcagttcctg tcccttctct taatttaatc ttttttatgt gccgtgttat tgtattaggt 540  
gtcatttcca ttatttatat tagtttagcc aaaggataag tgtcctatgg ggatgggtcca 600  
ctgtcactgt ttctctgctg ttgcaaatac atggataaca catttgattc tgttgtttt 660  
ccataataaa actttaaaat aaaatgcaga cagtta 696

<210> 32

<211> 3338

<212> DNA

<213> NM\_001455.2| Homo sapiens forkhead box O3A (FOXO3A), transcript variant 1, mRNA

<400> 32  
gcgcgaggcc gtcgattcgc tcgcggctcc atcgcggcct ggccgggggg cggtgtctgc 60  
tgcgccaggt tcgctggccg cacgtcttca ggtcctcctg ttcctgggag gcgggcgcgg 120  
caggactggg aggtggcggc agcgggcgag gactcgccga ggacggggct ccggccccggg 180  
ataaccaact ctcttctctt cttctttggt gcttccccag gcggcggcgg cggcgccccg 240  
gagccggagc cttecgggcg tccacgtccc tccccgctg cccccgccc cggcgcgaga 300  
ggagagcgcg agagccccag ccgcgggcgg gcgggcggcg aagatggcag aggcaccggc 360

ttccccggcc ccgctctctc cgctcgaagt ggagctggac ccggagttcg agccccagag	420
ccgtccgcga tcctgtacgt ggccccctgca aaggccggag ctccaagcga gccctgccaa	480
gccctcgggg gagacggccg ccgactccat gatccccgag gaggaggacg atgaagacga	540
cgaggacggc gggggacggg ccggctcggc catggcgatc ggcggcggcg gcgggagcgg	600
cacgctgggc tccgggctgc tccttgagga ctcggccccg gtgctggcac ccggagggca	660
agacccccggg tctggggccag ccaccgcggc gggcgggctg agcgggggta cacaggcgct	720
gctgcagcct cagcaaccgc tgccaccgcc gcagccgggg gcggctgggg gctccgggca	780
gccgaggaaa tgttcgtcgc ggcggaacgc ctggggaaac ctgtcctacg cggacctgat	840
caccgcgcgc atcgagagct ccccggaaca acggctcact ctgtcccaga tctacgagt	900
gatggtgcgt tgcgtgccct acttcaagga taagggcgac agcaacagct ctgccggctg	960
gaagaactcc atccggcaca acctgtcact gcatagtcga ttcattgcggg tccagaatga	1020
gggaactggc aagagctctt ggtggatcat caacctgat ggggggaaga gcggaaaagc	1080
cccccggcgg cgggctgtct ccatggacaa tagcaacaag tataccaaga gccgtggccg	1140
cgcagccaag aagaaggcag ccctgcagac agccccgaa tcagctgacg acagtccctc	1200
ccagctctcc aagtggcctg gcagccccac gtcacgcagc agtgatgagc tggatgcgtg	1260
gacggacttc cgttcacgca ccaattctaa cgccagcaca gtcagtggcc gcctgtcgcc	1320
catcatggca agcacagagt tggatgaagt ccaggacgat gatgcgcctc tctcgcccat	1380
gctctacagc agctcagcca gcctgtcacc ttcagtaagc aagccgtgca cgggtggaact	1440
gccacggctg actgatatgg caggcaccat gaattctgaat gatgggctga ctgaaaacct	1500
catggacgac ctgctggata acatcacgct cccgccatcc cagccatcgc ccaactggggg	1560
actcatgcag cggagctcta gcttcccgta taccaccaag ggctcggggc tgggctcccc	1620
aaccagctcc tttaacagca cgggtgttcg accctcatct ctgaactccc tacgccagtc	1680
tcccatgcag accatccaag agaacaagcc agctaccttc tcttccatgt cacactatgg	1740
taaccagaca ctccaggacc tgctcacttc ggactcactt agccacagcg atgtcatgat	1800
gacacagtcg gaccccttga tgtctcaggc cagcaccgct gtgtctgccc agaattcccg	1860
ccggaacgtg atgcttcgca atgatccgat gatgtccttt gctgcccagc ctaaccaggg	1920
aagtttggtc aatcagaact tgctccacca ccagcaccaa acccagggcg ctcttggtgg	1980
cagccgtgcc ttgtcgaatt ctgtcagcaa catgggcttg agtgagtcca gcagccttg	2040
gtcagccaaa caccagcagc agtctcctgt cagccagtct atgcaaacc tctcggactc	2100
tctctcaggc tcctccttgt actcaactag tgcaaacctg cccgtcatgg gccatgagaa	2160
gttccccagc gacttggacc tggacatgtt caatgggagc ttggaatgtg acatggagtc	2220
cattatccgt agtgaactca tggatgctga tgggttggtt tttaactttg attccctcat	2280
ctccacacag aatgttggtg gtttgaacgt ggggaacttc actggtgcta agcaggcctc	2340
atctcagagc tgggtgccag gctgaaggat cactgaggaa ggggaagtgg gcaaagcaga	2400

```

ccctcaaact gacacaagac ctacagagaa aaccctttgc caaatctgct ctcagcaagt 2460
ggacagtgat accgtttaca gcttaacacc tttgtgaatc ccacgccatt ttcctaaccc 2520
agcagagact gttaatggcc ccttaccctg ggtgaagcac ttacccttgg aacagaactc 2580
taaaaagtat gcaaaatctt ccttgtagag ggtggtgagc cgcctgccag tggaggacag 2640
caccctcag caccaccac cctcattcag agcacaccgt gagccccctg cggccattct 2700
gtggtgtttt aatattgcga tggtttatgg gacgttttaa gtgttggtct tgtgtttgtt 2760
ttcctttgac tttctgagtt tttcacatgc attaacttgc ggtatttttc tgttaaaatg 2820
ttaaccgtcc tttccctagc aaatttaaaa acagaaagaa aatgttgtag cagttaccat 2880
tccgggttcg agcatcacia gcttttgagc gcatggaact ccataaacta acaaattaca 2940
taaactaaag ggggattttc tttcttcttt tgtttggtag aaaattatcc ttttctaaaa 3000
actgaacaat ggcacaattg tttgctatgt gcacccgtcc aggacagaac cgtgcatagg 3060
caaaaggagt ggagcacagc gtccggccca gtgtgtttcc ggttctgagt caggggtgatc 3120
tgtggacggg accccagcac caagtctacg ggtgccagat cagtagggcc tgtgatttcc 3180
tgtcagtgtc ctcagctaata gtgaacagtg ttggtctgct ggtagaaac tagaatattg 3240
atattttcag gaaagaaatc agctcagctc tccactcatt gccaaatgtc actaaagggt 3300
ttagttttta ggagaaagaa aaggaaaaaa aaaaaaaa 3338

```

<210> 33

<211> 2646

<212> DNA

<213> NM\_152873.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 4, mRNA

```

<400> 33
cctacccgcg cgcaggccaa gttgctgaat caatggagcc ctccccaacc cgggcgttcc 60
ccagcgaggc ttccttccca tctcctgac caccggggct tttcgtgagc tcgtctctga 120
tctcgcgcaa gagtgcaca caggtgttca aagacgcttc tggggagtga ggggaagcgtt 180
ttacgagtga cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggg 240
cagccctggc tgcccaggcg gagctgcctc ttctcccgcg ggttggtgga cccgctcagt 300
acggagttgg ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg 360
gaccctccta cctctgggtc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc 420
ccaagtgact gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga 480
gactcagaac ttggaaggcc tgcatcatga tggccaattc tgccataagc cctgtcctcc 540
aggtgaaagg aaagctaggg actgcacagt caatggggat gaaccagact gcgtgccctg 600
ccaagaaggg aaggagtaca cagacaaagc ccatttttct tccaaatgca gaagatgtag 660

```

atttgtgtgat gaaggacatg atgtgaacat ggaatcatca aggaatgcac actcaccagc	720
aacaccaagt gcaaagagga aggatccaga tctaacttgg ggtggctttg tcttcttctt	780
ttgccaatc cactaattgt ttgggtgaag agaaagggaag tacagaaaac atgcagaaaag	840
cacagaaaagg aaaaccaagg ttctcatgaa tctccaacct taaatcctga aacagtggca	900
ataaatttat ctgatgttga cttgagtaaa tatatcacca ctattgctgg agtcatgaca	960
ctaagtcaag ttaaaggctt tgttcgaaag aatggtgtca atgaagccaa aatagatgag	1020
atcaagaatg acaatgtcca agacacagca gaacagaaaag ttcaactgct tcgtaattgg	1080
catcaacttc atggaaagaa agaagcgtat gacacattga ttaaagatct caaaaaagcc	1140
aatctttgta ctcttgcaag gaaaattcag actatcatcc tcaaggacat tactagtac	1200
tcagaaaatt caaacttcag aaatgaaatc caaagcttgg tctagagtga aaaacaacaa	1260
attcagttct gagtatatgc aattagtgtt tgaaaagatt cttaatagct ggctgtaaat	1320
actgcttgggt tttttactgg gtacatttta tcatttatta gcgctgaaga gccaacatat	1380
ttgtagattt ttaatatctc atgattctgc ctccaaggat gtttaaaatc tagttgggaa	1440
aacaaacttc atcaagagta aatgcagtgg catgctaagt acccaaatag gagtgtatgc	1500
agaggatgaa agattaagat tatgctctgg catctaacat atgattctgt agtatgaatg	1560
taatcagtgt atgttagtac aaatgtctat ccacaggcta accccactct atgaatcaat	1620
agaagaagct atgacctttt gctgaaatat cagttactga acaggcaggc cactttgcct	1680
ctaaattacc tctgataatt ctgagattt taccatattt ctaaaacttg tttataactc	1740
tgagaagatc atatttatgt aaagtatatg tatttgagtg cagaatttaa ataaggctct	1800
acctcaaaga cctttgcaca gtttattgggt gtcatattat acaatatttc aattgtgaat	1860
tcacatagaa aacattaaat tataatgttt gactattata tatgtgtatg cttttactg	1920
gctcaaaact acctacttct ttctcaggca tcaaaagcat tttagcagg agagtattac	1980
tagagctttg ccacctctcc atttttgcct tgggtgctcat cttaatggcc taatgcaccc	2040
ccaaacatgg aaatatcacc aaaaaatact taatagtcca ccaaaggca agactgcct	2100
tagaaattct agcctggttt ggagatacta actgctctca gagaaagtag ctttgtgaca	2160
tgcatgaac ccatgtttgc aatcaaagat gataaaatag attcttattt ttccccacc	2220
cccgaaaatg ttcaataatg tcccatgtaa aacctgctac aaatggcagc ttatacatag	2280
caatggtaaa atcatcatct ggatttagga attgctcttg tcatacccc aagtttctaa	2340
gatttaagat tctccttact actatcctac gtttaaatat ctttgaaagt ttgtattaaa	2400
tgtgaatttt aagaaataat atttatattt ctgtaaatgt aaactgtgaa gatagttata	2460
aactgaagca gatacctgga accacctaaa gaacttccat ttatggagga ttttttgcc	2520
ccttgtgttt ggaattataa aatataggta aaagtacgta attaaataat gtttttgga	2580
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	2640

aaaaaa

2646

<210> 34

<211> 817

<212> DNA

<213> NM\_002038.2| Homo sapiens interferon, alpha-inducible protein (clone IFI-6-16) (G1P3), transcript variant 1, mRNA

<400> 34  
gaaccgttta ctcgctgctg tgcccatcta tcagcaggct ccgggctgaa gattgcttct 60  
cttctctcct ccaaggctcta gtgacggagc ccgcgcgcgg cgccaccatg cggcagaagg 120  
cggtatcgct ttctttgtgc tacctgctgc tcttcaactg cagtgggggtg gaggcaggta 180  
agaaaaagtg ctcggagagc tcggacagcg gctccggggt ctggaaggcc ctgaccttca 240  
tgggcgtcgg agggaggactc gcagtcgccc ggctgcccgc gctgggcttc accggcgccg 300  
gcatcgccgc caactcgggtg gctgcctcgc tgatgagctg gtctgcgata ctgaatgggg 360  
gcggcgtgcc cgccgggggg ctagtgggca cgctgcagag cctcgggggt ggtggcagca 420  
gcgtcgtcat aggtaatat ggtgccctga tgggctacgc caccacaag tatctcgata 480  
gtgaggagga tgaggagtag ccagcagctc ccagaacctc ttcttccttc ttggcctaac 540  
tcttccagtt aggatctaga actttgcctt tttttttttt tttttttttt tttgagatgg 600  
gttctcacta tattgtccag gctagagtgc agtggctatt cacagatgcg aacatagtag 660  
actgcagcct ccaactccta gcctcaagtg atcctcctgt ctcaacctcc caagtaggat 720  
tacaagcatg cgccgacgat gccagaatc cagaactttg tctatcactc tccccaacaa 780  
cctagatgtg aaaacagaat aaacttcacc cagaaaa 817

<210> 35

<211> 1172

<212> DNA

<213> NM\_001565.1| Homo sapiens chemokine (C-X-C motif) ligand 10 (CXCL10), mRNA

<400> 35  
gagacattcc tcaattgctt agacatattc tgagcctaca gcagaggaac ctccagtctc 60  
agcaccatga atcaaactgc gattctgatt tgctgcctta tctttctgac tctaagtggc 120  
attcaaggag tacctctctc tagaacgta cgctgtacct gcatcagcat tagtaatcaa 180  
cctgttaatc caaggctttt agaaaaactt gaaattattc ctgcaagcca attttgtcca 240  
cgtgttgaga tcattgctac aatgaaaaag aagggtgaga agagatgtct gaatccagaa 300  
tcgaaggcca tcaagaattt actgaaagca gttagcaagg aaatgtctaa aagatctcct 360

taaaaccaga ggggagcaaa atcgatgcag tgcttccaag gatggaccac acagaggctg	420
cctctcccat cacttcccta catggagtat atgtcaagcc ataattgttc ttagtttgca	480
gttactactaa aaggtgacca atgatggtca ccaaatacagc tgctactact cctgtaggaa	540
ggttaatggt catcatccta agctattcag taataactct accctggcac tataatgtaa	600
gctctactga ggtgctatgt tcttagtgga tgttctgacc ctgcttcaaa tatttccttc	660
acctttccca tcttccaagg gtactaagga atctttctgc tttgggggtt atcagaattc	720
tcagaatctc aaataactaa aaggatgca atcaaactcg ctttttaaag aatgctcttt	780
acttcatgga cttccactgc catcctccca aggggcccaa attctttcag tggctaccta	840
catacaattc caaacacata caggaaggta gaaatatctg aaaatgtatg tgtaagtatt	900
ctratttaat gaaagactgt acaaagtata agtcttagat gtatatattt cctatatgtt	960
tttcagtgt catggaataa catgtaatta agtactatgt atcaatgagt aacaggaaaa	1020
ttttaaaaat acagatagat atatgctctg catgttacat aagataaatg tgctgaatgg	1080
ttttcaaata aaaatgaggt actctcctgg aaatattaag aaagactatc taaatgttga	1140
aagatcaaaa ggttaataaa gtaattataa ct	1172

<210> 36

<211> 396

<212> DNA

<213> NM\_005950.1| Homo sapiens metallothionein 1G (MT1G), mRNA

<400> 36	
actccgcctt ccacgtgcac ccactgcctc ttcccttctc gcttgggaac tctagtctcg	60
cctcggggttg caatggaccc caactgctcc tgtgccgctg gtgtctcctg cacctgcgcc	120
agctcctgca agtgcaaaga gtgcaaatgc acctcctgca agaagagctg ctgctcctgc	180
tgccctgtgg gctgtgccaa gtgtgcccaa ggctgcatct gcaaaggggc atcggagaag	240
tgcagctgct gcgcctgatg tcgggacagc cctgctccca agtacaata gagtgacccg	300
taaaatctag gattttttgt tttttgctac aatcttgacc cttttgctac attccctttt	360
ttctgtgaaa tatgtgaata ataattaaac acttag	396

<210> 37

<211> 2755

<212> DNA

<213> NM\_000043.3| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 1, mRNA



<400> 37  
cctacccgcg cgcaggccaa gttgctgaat caatggagcc ctccccaacc cgggcgttcc 60  
ccagcgaggc ttccttccca tcctcctgac caccggggct tttcgtgagc tcgtctctga 120  
tctcgcgcaa gagtgcaca caggtgttca aagacgcttc tggggagtga gggaaagcgg 180  
ttacgagtga cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggc 240  
cagccctggc tgcccaggcg gagctgcctc ttctcccgcg ggttggtgga cccgctcagt 300  
acggagttgg ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg 360  
gacctccta cctctggttc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc 420  
ccaagtgact gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga 480  
gactcagaac ttggaaggcc tgcattcatga tggccaattc tgccataagc cctgtcctcc 540  
aggtgaaagg aaagctaggg actgcacagt caatggggat gaaccagact gcgtgccctg 600  
ccaagaaggg aaggagtaca cagacaaagc ccatttttct tccaaatgca gaagatgtag 660  
attgtgtgat gaaggacatg gcttagaagt ggaaataaac tgcacccgga cccagaatac 720  
caagtgcaga tgtaaacc aaactttttt taactctact gtatgtgaac actgtgaccc 780  
ttgcaccaa tgtgaacatg gaatcatcaa ggaatgcaca ctcaccagca acaccaagt 840  
caaagaggaa ggatccagat ctaacttggg gtggctttgt cttcttcttt tgccaattcc 900  
actaattgtt tgggtgaaga gaaaggaagt acagaaaaca tgcagaaagc acagaaagga 960  
aaaccaaggt tctcatgaat ctccaacctt aaatcctgaa acagtggcaa taaatttatc 1020  
tgatgttgac ttgagtaa atatcaccac tattgctgga gtcattgacac taagtcaagt 1080  
taaaggcttt gttcgaaaga atggtgtcaa tgaagccaaa atagatgaga tcaagaatga 1140  
caatgtccaa gacacagcag aacagaaagt tcaactgctt cgtaattggc atcaacttca 1200  
tggaagaaa gaagcgtatg acacattgat taaagatctc aaaaaagcca atctttgtac 1260  
tcttgagag aaaattcaga ctatcatcct caaggacatt actagtgact cagaaaattc 1320  
aaacttcaga aatgaaatcc aaagcttggg ctagagtga aaacaacaaa ttcagttctg 1380  
agtatatgca attagtgttt gaaaagattc ttaatagctg gctgtaaata ctgcttgggt 1440  
ttttactggg tacattttat catttattag cgctgaagag ccaacatatt tgtagatttt 1500  
taatattctca tgattctgcc tccaaggatg tttaaaatct agttgggaaa acaaacttca 1560  
tcaagagtaa atgcagtggc atgctaagta cccaaatagg agtgatgca gaggatgaaa 1620  
gattaagatt atgctctggc atctaacata tgattctgta gtatgaatgt aatcagtgt 1680  
tgtagtaga aatgtctatc cacaggctaa cccactcta tgaatcaata gaagaagcta 1740  
tgaccttttg ctgaaatatt agttactgaa caggcaggcc actttgcctc taaattacct 1800  
ctgataattc tagagatttt accatatttc taaactttgt ttataactct gagaagatca 1860  
tatttatgta aagtatatgt atttgagtgc agaatttaaa taaggctcta cctcaaagac 1920  
ctttgcacag tttattgggtg tcatattata caatatttca attgtgaatt cacatagaaa 1980

acattaaatt ataatgtttg actattatat atgtgtatgc attttactgg ctcaaaacta	2040
cctacttctt tctcagggcat caaaagcatt ttgagcagga gagtattact agagctttgc	2100
cacctctcca tttttgcctt ggtgctcatc ttaatggcct aatgcacccc caaacatgga	2160
aatatcacca aaaaatactt aatagtccac caaaaggcaa gactgccctt agaaattcta	2220
gcctggtttg gagatactaa ctgctctcag agaaagtagc tttgtgacat gtcatgaacc	2280
catgtttgca atcaaagatg ataaaataga ttcttatttt tccccacccc ccgaaaatgt	2340
tcaataatgt cccatgtaaa acctgctaca aatggcagct tatacatagc aatggtaaaa	2400
tcatcatctg gatttaggaa ttgctcttgt catacccca agtttctaag atttaagatt	2460
ctccttacta ctatcctacg tttaaatatac tttgaaagtt tgtattaaat gtgaatttta	2520
agaaataata tttatatattc tgtaaagtga aactgtgaag atagttataa actgaagcag	2580
atacctggaa ccacctaaag aacttccatt tatggaggat ttttttgccc cttgtgtttg	2640
gaattataaa atataggtaa aagtacgtaa ttaaataatg tttttggtaa aaaaaaaaaa	2700
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa	2755

<210> 38

<211> 1600

<212> DNA

<213> NM\_001953.2| Homo sapiens endothelial cell growth factor 1 (platelet-derived) (ECGF1), mRNA

<400> 38	
gccccgccgc cggcagtgga ccgctgtgcg cgaaccctga accctacggt cccgacccgc	60
gggcgaggcc gggtagcttg gctgggatcc ggagcaagcg ggcgagggca gcgccctaag	120
caggccccga gcgatggcag ccttgatgac cccgggaacc ggggccccac ccgcgcctgg	180
tgacttctcc ggggaagggg gccagggact tcccgaccct tcgccagagc ccaagcagct	240
cccggagctg atccgcatga agcgagacgg aggccgcctg agcgaagcgg acatcagggg	300
cttcgtggcc gctgtggtga atgggagcgc gcagggcgca cagatcgggg ccatgctgat	360
ggccatccga cttcggggca tggatctgga ggagacctcg gtgctgaccc aggccctggc	420
tcagtcggga cagcagctgg agtggccaga ggcctggcgc cagcagcttg tggacaagca	480
ttccacaggg ggtgtgggtg acaaggtcag cctggtcctc gcacctgccc tggcggcatg	540
tggctgcaag gtgccaatga tcagcggacg tggctctgggg cacacaggag gcaccttgga	600
taagctggag tctattcctg gattcaatgt catccagagc ccagagcaga tgcaagtgct	660
gctggaccag gcgggctgct gtatcgtggg tcagagttag cagctgggtc ctgcggacgg	720
aatcctatat gcagccagag atgtgacagc caccgtggac agcctgccac tcatcacagc	780
ctccattctc agtaagaaac tcgtggaggg gctgtccgct ctggtggtgg acgttaagtt	840

cggagggggcc gccgtcttcc ccaaccagga gcaggcccgg gagctggcaa agacgctggt	900
tggcgtggga gccagcctag ggcttcgggt cgcggcagcg ctgaccgcca tggacaagcc	960
cctgggtcgc tgcgtgggcc acgccctgga ggtggaggag gcgctgctct gcatggacgg	1020
cgcaggcccc ccagacttaa gggacctggt caccacgctc gggggcgccc tgctctggct	1080
cagcggacac gcggggactc aggctcaggg cgctgcccgg gtggccgagg cgctggacga	1140
cggtcgggcc ctgtggcggct tcgagcggat gctggcggcg cagggcgagg atccccggtct	1200
ggccccgagcc ctgtgctcgg gaagtcccg cagaacgccc cagctgctgc ctgcgcggcg	1260
ggagcaggag gagctgctgg cggccgcaga tggcaccgtg gagctggtcc gggcgctgcc	1320
gctggcgctg gtgctgcacg agctcggggc cgggcgcagc gcgctggggg agccgctccg	1380
cctggggggtg ggcgcagagc tgctgggtcga cgtgggtcag aggctgcgcc gtgggacccc	1440
ctggctccgc gtgcaccggg acggccccgc gctcagcggc ccgcagagcc gcgccctgca	1500
ggaggcgctc gtactctccg accgcgcgcc attcgccgcc ccctcgccct tcgcagagct	1560
cgttctgccg ccgcagcaat aaagctcctt tgccgcgaaa	1600

<210> 39

<211> 931

<212> DNA

<213> NM\_005138.1| Homo sapiens SCO cytochrome oxidase deficient homolog 2 (yeast) (SCO2), nuclear gene encoding mitochondrial protein, mRNA

<400> 39	
gcagagccca gggagctgga ggtcggcgct tcctctcgtg cttggtccac tgacgcgcgg	60
ccccgcgcgc aggagcatca gatccatgct gctgctgact cggagcccca cagcttggca	120
caggctctct cagctcaagc ctccggctct ccctgggacc ctgggaggcc agggccctgca	180
tctgaggtcc tggcttttgt caaggcaggg ccctgcagag acagggtggg agggccagcc	240
ccagggccct gggcttcgaa cccggctgct gatcacaggc ctgttcgggg ctggactcgg	300
tggggcctgg ctggccctga gggctgagaa ggagaggctg cagcagcaaa agcgaacaga	360
agccctgcgc caggcagctg tgggccaggg cgacttccac ctgctggatc acagaggccg	420
ggctcgctgc aaggctgact tccggggcca gtgggtgctg atgtactttg gcttactca	480
ctgccctgac atctgcccag acgagctgga gaagctggtg cagggtggtg ggcagctgga	540
agcagagcct ggtttgccct cagtgcagcc tgtcttcac actgtggacc ccgagcggga	600
cgacgttgaa gccatggccc gctacgtcca ggacttccac ccaagactgt tgggtctgac	660
cggctccacc aaacagggtt gccaggctag tcacagttac gcggtgtact acaatgccgg	720
ccccaaaggat gaggaccagg actacatcgt ggaccactcc attgccatct acctgctcaa	780
ccctgacggc ctcttcacgg attactacgg ccggagcaga tcggctgagc agatctcaga	840

cagtgtgcgg cggcacatgg cggctttccg cagtgtcctg tcttgagcca ctgcagtctg 900  
ggcccatca ttaaacgggc tgcgtttaaa a 931

<210> 40

<211> 1216

<212> DNA

<213> NM\_006419.1| Homo sapiens chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant) (CXCL13), mRNA

<400> 40  
ttcggcactt gggagaagat gtttgaaaaa actgactctg ctaatgagcc tggactcaga 60  
gctcaagtct gaactctacc tccagacaga atgaagtcca tctcgacatc tctgcttctc 120  
atgctgctgg tcagcagcct ctctccagtc caaggtgttc tggaggtcta ttacacaagc 180  
ttgaggtgta gatgtgtcca agagagctca gtctttatcc ctagacgctt cattgatcga 240  
attcaaatct tgccccgtgg gaatggttgt ccaagaaaag aaatcatagt ctggaagaag 300  
aacaagtcaa ttgtgtgtgt ggaccctcaa gctgaatgga taaaagaat gatggaagta 360  
ttgagaaaaa gaagttcttc aactctacca gttccagtgt ttaagagaaa gattccctga 420  
tgctgatatt tccactaaga acacctgcat tcttccctta tccctgctct ggattttagt 480  
tttgtgctta gttaaactct ttccagggag aaagaacttc cccatacaaa taaggcatga 540  
ggactatgtg aaaaataacc ttgcaggagc tgatggggca aactcaagct tcttcactca 600  
cagcacccta tatacacttg gagtttgcat tcttattcat cagggaggaa agtttctttg 660  
aaaatagtta ttcagttata agtaatacag gattattttg attatatact tgttgtttaa 720  
tgtttaaaat ttcttagaaa acaatggaat gagaatttaa gcctcaaatt tgaacatgtg 780  
gcttgaatta agaagaaaat tatggcatat attaaaagca ggcttctatg aaagactcaa 840  
aaagctgcct gggaggcaga tggaacttga gcctgtcaag aggcaaagga atccatgtag 900  
tagatatcct ctgcttaaaa actcactacg gaggagaatt aagtcctact tttaaagaat 960  
ttctttataa aatttactgt ctaagattaa tagcattcga agatccccag acttcataga 1020  
atactcaggg aaagcattta aagggtgatg tacacatgta tcctttcaca catttgcctt 1080  
gacaaacttc tttcactcac atctttttca ctgacttttt ttgtgggggc ggggccgggg 1140  
ggactctggt atctaattct ttaatgattc ctataaatct aatgacattc aataaagttg 1200  
agcaaacatt ttactt 1216

<210> 41

<211> 738

<212> DNA

<213> NM\_006433.2| Homo sapiens granulysin (GNLY), transcript variant NKG5, mRNA

```

<400> 41
gtatctgtgg taaacccagt gacacggggg agatgacata caaaaagggc aggacctgag 60
aaagattaag ctgcaggctc cctgcccata aaacaggggtg tgaaaggcat ctcagcggct 120
gccccaccat ggctacctgg gccctcctgc tccttgacgc catgctcctg ggcaacccag 180
gtctggctctt ctctcgtctg agccctgagt actacgacct ggcaagagcc cacctgcgtg 240
atgaggagaa atcctgcccc tgccctggccc aggagggccc ccagggtgac ctgttgacca 300
aaacacagga gctgggccgt gactacagga cctgtctgac gatagtccaa aaactgaaga 360
agatgggtgga taagcccacc cagagaagtg tttccaatgc tgcgacccgg gtgtgtagga 420
cggggaggtc acgatggcgc gacgtctgca gaaatttcat gaggaggat cagtctagag 480
ttaccacagg cctcgtggcc ggagaaactg cccagcagat ctgtgaggac ctcaggttgt 540
gtataccttc tacaggctcc ctctgagccc tctcaccttg tcctgtggaa gaagcacagg 600
ctcctgtcct cagatcccgg gaacctcagc aacctctgcc ggctcctcgc ttcctcgatc 660
cagaatccac tctccagtct cctccctctg actccctctg ctgtcctccc ctctcacgag 720
aataaagtgt caagcaag 738

```

<210> 42

<211> 1579

<212> DNA

<213> NM\_001767.2| Homo sapiens CD2 antigen (p50), sheep red blood cell receptor (CD2), mRNA

```

<400> 42
accaaccctt aagatgagct ttccatgtaa atttgtagcc agcttccttc tgattttcaa 60
tgtttcttcc aaaggtgcag tctccaaaga gattacgaat gccttggaag cctgggggtgc 120
cttgggtcag gacatcaact tggacattcc tagttttcaa atgagtgatg atattgacga 180
tataaaatgg gaaaaaactt cagacaagaa aaagattgca caattcagaa aagagaaaga 240
gactttcaag gaaaaagata catataagct atttaaaaat ggaactctga aaattaagca 300
tctgaagacc gatgatcagg atatctacaa ggtatcaata tatgatacaa aaggaaaaaa 360
tgtgttgga aaaaatatttg atttgaagat tcaagagagg gtctcaaaac caaagatctc 420
ctggacttgt atcaacacaa ccctgacctg tgaggtaatg aatggaactg accccgaatt 480
aaacctgtat caagatggga aacatctaaa actttctcag agggatcatc cacacaagtg 540
gaccaccagc ctgagtgcag aattcaagtg cacagcaggg aacaaagtca gcaaggaatc 600
cagtgtcgag cctgtcagct gtccagagaa aggtctggac atctatctca tcattggcat 660

```

atgtggagga ggcagcctct tgatggtctt tgtggcactg ctcgttttct atatcaccaa	720
aaggaaaaaa cagaggagtc ggagaaatga tgaggagctg gagacaagag cccacagagt	780
agctactgaa gaaaggggcc ggaagcccca acaaattcca gttcaaccc ctcagaatcc	840
agcaacttcc caacatcctc ctccaccacc tggtcacgt tcccaggcac ctagtcacgt	900
tccccgcct cctggacacc gtgttcagca ccagcctcag aagaggcctc ctgctccgtc	960
gggcacacaa gttcaccagc agaaaggccc gccctcccc agacctcgag ttcagccaaa	1020
acctcccat ggggcagcag aaaactcatt gtcccccttc tctaattaaa aaagatagaa	1080
actgtctttt tcaataaaaa gcactgtgga tttctgccct cctgatgtgc atatccgtac	1140
ttccatgagg tgttttctgt gtgcagaaca ttgtcacctc ctgaggctgt gggccacagc	1200
cacctctgca tcttcgaact cagccatgtg gtcaacatct ggagtttttg gtctcctcag	1260
agagctccat cacaccagta aggagaagca atataagtgt gattgcaaga atggtagagg	1320
accgagcaca gaaatcttag agatttcttg tccccctcga ggtcatgtgt agatgcgata	1380
aatcaagtga ttggtgtgcc tgggtctcac tacaagcagc ctatctgctt aagagactct	1440
ggagtttctt atgtgccctg gtggacactt gccaccatc ctgtgagtaa aagtgaaata	1500
aaagctttga ctagaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1560
aaaaaaaaaa aaaaaaaaaa	1579

<210> 43

<211> 3738

<212> DNA

<213> NM\_006275.4| Homo sapiens splicing factor, arginine/serine-rich 6 (SFRS6), mRNA

<400> 43	
ctggcgcgcg cgcgcgccat tgtgtggctg gactcggccg cccctgtggt gtgaggcgcg	60
tggtcgggct cttgccgtcc ccgcacccgc accgcgggta ctggcttgcg gtccgccgtt	120
cgacaaccag cccttgggtc cccgcccgc accgacatgc cgcgcgtcta cataggacgc	180
ctgagctaca acgtccggga gaaggacatc cagcgtttt tcagtggcta tggccgcctc	240
ctcgaagtag acctcaaaaa tgggtacggc ttcgtggagt tgcaggactc ccgcgacgcc	300
gacgacgccg tttagcagct gaacggcaag gagctctgcg gcgagcgcgt gatcgtagag	360
cacgcccggg gcccgcgctc cgatcgcgac ggctacagct acggaagccg cagtgggtgga	420
gggtggataca gcagtcggag aacatctggc agagacaaat acggaccacc tggtcgtaca	480
gaatacaggc ttattgtaga aaatctttct agtcggtgca gttggcaaga tttaaaggat	540
tttatgcgac aagcaggtag agtaacctat gcggatgcc acaaggaacg aacaaatgag	600
gggtgaattg agtttcgctc ctactctgac atgaagcgtg ctttggacaa actggatggc	660

acagaaataa	atggcagaaa	tattaggctt	attgaagata	agccacgcac	aagccatagg	720
cgatcttact	ctggaagcag	atccaggtct	cgatctagaa	gacggtcacg	aagtaggagt	780
cgcaggagca	gccgcagtag	atctcgaagt	atctcaaaaa	gtcgcctccc	ttccaggctc	840
cggagcaaa	gtcgatcacg	ttctcgatca	aaaggcagga	aatctagatc	aaagagcaaa	900
tctaagccca	agtctgatcg	gggctcccat	tcacattctc	gaagcagatc	taaggatgag	960
tatgagaaat	ctcgaagcag	gtctcgggtc	cgatccccc	aagaaaatgg	aaaggggtgat	1020
ataaagtcaa	aatccagatc	aaggagccag	tcccgttcca	attcgccgct	acctgttcca	1080
ccctcaaagg	cccgtttctgt	gtcccttcca	ccaaaaagag	ctacttcaag	atcccgttct	1140
agatctcgct	caaagtcaag	atcaagggtc	aggctcgagtt	ccagagatta	actcagaact	1200
ccttgtttgc	acattattat	ggaacacttt	cctacttagg	cagttactct	tccatgttta	1260
tacttggcct	cttctgcaag	aggaatctct	tgaaaacagg	ggcacacaga	aatttgattt	1320
gtggccaaat	tggatgaaaa	agatgaggct	ctaaggaaat	ggtggcatga	agaccctctc	1380
ccttctttgt	agaattaaga	taactttgat	tttatagctt	ttgagctaac	gtaacttttg	1440
taaagattaa	gctcatttag	tgttgttttt	tttttttttt	tttttttttt	tttttagtat	1500
ttcagcagga	tctgctggca	gggttttttt	gttttatttg	tttgcttatt	tttaaattaa	1560
ctgttttgag	ctttgaatac	ttaaggcttt	agagggagaa	cccaattttc	aattatgttg	1620
gctttttata	aagcttgagt	tatgtaagat	ttaaataaaa	gtttgctacc	aagatgattg	1680
ccttattgaa	taggtcacta	ttaaattcct	ttaaatgttg	atatctgcc	tttgtggaaa	1740
caacgtaaat	tctacttaag	tgtaaacaag	gcaagcctca	gaccagcaat	aaattactca	1800
gtttggataa	cattattttg	tgcagttaat	caaatttgcc	aaagtcttta	tctgccccct	1860
taacaagttg	agtaaaaata	aaaggatatt	tttagtcaat	gtgttccatg	attttgctta	1920
aattaatact	tttaagtaat	ggaacttttt	tcaaaggcaa	atttaaacta	tttaagaaat	1980
agctccta	acttgggatc	ttgttttag	aatccacttt	ctggaagtcc	tcagcataat	2040
tagtgttgag	agtggttcag	ttgtctttta	tgtttgtcat	gtggaaatgg	aagtagcctc	2100
tttttgttct	gaaattgagt	ttattcaaag	tgtaaaagca	catactgcat	tttctgctga	2160
aagatcatta	tgtttaacag	gcacttaatc	tcagtaaagt	cagttgccag	ttaagttcca	2220
cccagtagtc	agtccccctt	gtagttagtg	ggattatttg	ataattgggt	agatcatact	2280
tgtaaatttt	aatgctttgt	gtaattgggt	tgaaaaacag	tgaaatgggt	aaacgcaaaa	2340
cttttgtact	ttattacgag	taaagtgtaa	tgagtactgt	ggaaaccaa	tttgaatact	2400
gcaaatttgt	aggagttagt	agggttagca	ttagtccata	catccataag	cctgatgagt	2460
tgaaattgca	gtttgagaag	tgaattaacc	ttacatccct	ttgttcagat	accttaaaag	2520
ttactttatt	taaaagcatt	tattaatctt	agtctgaaat	caaaatatag	attaattggc	2580
tcagctttta	tacctttcta	ggagggtgtc	caatgtaggg	taccaaggg	tggattgtga	2640
tggggcatgg	tcgtacactg	ctcattgtgc	cacagggtgtg	actggaaagc	atgataattct	2700

agggttggtt tgtagattca aataatccag aaatatacct aataagattg agtgaaaaat	2760
ttgagtcaaa tatctagggc attcacagag tagctgtgag ttcttggtta tgtgaaaaag	2820
gccttgtttt tcagaaattc ctgggtttcc tgttaaaaaa tcttaaagcc caaccttagg	2880
aatatagtgc cccaaaaggc ggatgcttct tccattatct tattttcttt gatactttat	2940
ttaattagat gtttataaag aaatgggttt atttttccag cataaacctc agaatttaag	3000
gaaagaaaat gatgtctgtt gttatagttc attgttttgc ctactcagca gaagtgatga	3060
ctcttaaaaa ttggctttga ccaaagttct cttgttttca gggaaagaac ataaaagctt	3120
tttgaactac agccttttta aaagagggat gggaggatat tacagtaaga aattaggctt	3180
tctaaaagta tgaaacatcc ttcaactggg ctctcttggt aataggacat catatggtaa	3240
tagactgggt tgactatatt gttagctgcc acagtaagca ggtcattgta taggtaaatg	3300
cctgcacca taattttcta gtaatagcca cgaccaatth attaacagtc agggcctatc	3360
cttgctgta gttctcagtc actggatgca caaaatcact gtgtaacatt ggctcacttg	3420
gtgagcatag ggttgactga taaaatgttt aattcccttg ctagcttggt agaagaatga	3480
gttgatgaca tgctccatac cagtggctag atggagtatt aagggtggagc agaaaagaag	3540
tgagaacatc ttgattcccc ttctttttac ttgatgggtt ttatgaacat gccgtagtgc	3600
ctttatggcc agtttgagtc ctgcctactt tgacttttac gttcccattc ctgtgttacc	3660
accttctccc cgatttggtc acctattttg tgctttaaat ctcaataaaa tacttactga	3720
ggaaaaaaaa aaaaaaaaa	3738

<210> 44

<211> 2033

<212> DNA

<213> NM\_003212.1| Homo sapiens teratocarcinoma-derived growth factor 1 (TDGF1), mRNA

<400> 44

ggagaatccc cggaaaggct gagtctccag ctcaagggtca aaacgtccaa ggccgaaagc	60
cctccagttt cccctggacg cttgctcct gcttctgcta cgaccttctg gggaaaacga	120
atttctcatt ttcttcttaa attgccattt tcgcttttagg agatgaatgt tttcctttgg	180
ctgttttggc aatgactctg aattaaagcg atgctaacgc ctcttttccc cctaattggt	240
aaaagctatg gactgcagga agatggcccg cttctcttac agtgtgattt ggatcatggc	300
catttctaaa gtctttgaac tgggattagt tgccgggctg ggccatcagg aatttgctcg	360
tccatctcgg ggatacctgg ctttcagaga tgacagcatt tggccccagg aggagcctgc	420
aattcggcct cggctctccc agcgtgtgcc gcccatgggg atacagcaca gtaaggagct	480
aaacagaacc tgctgcctga atgggggaac ctgcatgctg gggtcctttt gtgcctgccc	540



tccctccttc	tacggacgga	actgtgagca	cgatgtgcmc	aaagagaact	gtgggtctgt	600
gccccatgac	acctggctgc	ccaagaagtg	ttccctgtgt	aaatgctggc	acggctcagct	660
ccgctgcttt	cctcaggcat	ttctacccgg	ctgtgatggc	cttgtgatgg	atgagcacct	720
cgtggcttcc	aggactccag	aactaccacc	gtctgcacgt	actaccactt	ttatgctagt	780
tggcatctgc	ctttctatac	aaagctacta	ttaatcgaca	ttgacctatt	tccagaaata	840
caatctttaga	tatcatgcaa	atttcatgac	cagtaaaggc	tgctgctaca	atgtcctaac	900
tgaaagatga	tcattttgtag	ttgccttaaa	ataatgaata	caatcttcaa	aatgggtctct	960
aacatcttct	tacagaacta	cttcttactt	ctttgccctg	ccctctccca	aaaaactact	1020
tctttttttca	aaagaaagtc	agccatatct	ccattgtgcc	taagtccagt	gtttcttttt	1080
tttttttttt	ttgagacgga	gtctcactct	gtcaccacgg	ctggactgca	atgacgcgat	1140
cttgggttcc	tgcaacctcc	gcatccgggg	ttcaagccat	tctcctgcct	aagcctccca	1200
agtaactggg	attacaggca	tgtgtcacca	tgcccagcta	atttttttgt	atttttagtag	1260
agatgggggt	ttcaccatat	tggccagtct	ggctctgaac	tctgaccttg	tgatccatcg	1320
atcagcctct	cgagtgtctga	gattacacac	gtgagcaact	gtgcaaggcc	tggtgtttct	1380
tgatacatgt	aattctacca	aggtcttctt	aatatgttct	tttaaagtat	tgaattatat	1440
gttcagatta	ttggagacta	attctaattg	ggaccttaga	atacagtttt	gagtagagtt	1500
gatcaaaatc	aattaaaata	gtctctttta	aaggaaagaa	aacatcttta	aggggaggaa	1560
ccagagtgtc	gaaggaatgg	aagtccatct	gcgtgtgtgc	agggagactg	ggtaggaaag	1620
aggaagcaaa	tagaagagag	aggttgaaaa	acaaaatggg	ttacttgatt	ggtgattagg	1680
tggtggtaga	gaagcaagta	aaaaggctaa	atggaagggc	aagtttccat	catctataga	1740
aagctatata	agacaagaac	tccccctttt	ttcccaaagg	cattataaaa	agaatgaagc	1800
ctccttagaa	aaaaaattat	acctcaatgt	ccccacaag	attgcttaat	aaattgtgtt	1860
tcctccaagc	tattcaattc	ttttaactgt	tgtagaagac	aaaatgttca	caatatattt	1920
agttgtaaac	caagtgatca	aactacatat	tgtaaagccc	atttttaaaa	tacattgtat	1980
atatgtgtat	gcacagtaaa	aatggaaact	atattgacct	aaaaaaaaaa	aaa	2033

<210> 45

<211> 367

<212> DNA

<213> NM\_005951.1| Homo sapiens metallothionein 1H (MT1H), mRNA

<400> 45

ctccagcttc	acctcggctt	gcaatggacc	ccaactgctc	ctgcgaggct	ggtggctcct	60
gcgcctgcgc	cggctcctgc	aagtgcaaaa	agtgcaaatg	cacctcctgc	aagaagagct	120

gctgctcctg ttgccccctg ggctgtgccca agtgtgcccc gggctgcatc tgcaaagggg	180
cgtcagagaa gtgcagctgc tgtgcctgat gtcgggacag ccctgctgtc agatgaaaac	240
agaatgacac gtaaaatccg aggttttttt tttctacaac tccgactcat ttgctacatt	300
cctttttttt tgtgaaatat gtgaataata attaaacact tagacttgaa aaaaaaaaaa	360
aaaaaaa	367

<210> 46

<211> 3052

<212> DNA

<213> NM\_000767.4| Homo sapiens cytochrome P450, family 2, subfamily B, polypeptide 6 (CYP2B6), mRNA

<400> 46	
caggaccatg gaactcagcg tcctcctctt ccttgcactc ctcacaggac tcttgctact	60
cctggttcag cgccacccta acacccatga ccgcctccca ccagggcccc gccctctgcc	120
cctttttgga aaccttctgc agatggatag aagaggccta ctcaaatcct ttctgaggtt	180
ccgagagaaa tatggggacg tcttcacggt acacctggga ccgaggcccg tggatcatgct	240
gtgtggagta gaggccatac gggaggccct tgtggacaag gctgaggcct tctctggccg	300
gggaaaaatc gccatggtcg acccattctt ccggggatat ggtgtgatct ttgccaatgg	360
aaaccgctgg aagggtgctt ggcgattctc tgtgaccact atgagggact tcgggatggg	420
aaagcggagt gtggaggagc ggattcagga ggaggctcag tgtctgatag aggagcttcg	480
gaaatccaag ggggccctca tggacccac cttcctcttc cagtccatta ccgccaacat	540
catctgctcc atcgtctttg gaaaacgatt ccaactacca gatcaagagt tcctgaagat	600
gctgaacttg ttctaccaga ctttttctact catcagctct gtattcggcc agctgtttga	660
gctcttctct ggcttcttga aatactttcc tggggcacac aggcaagttt acaaaaacct	720
gcaggaaatc aatgcttaca ttggccacag tgtggagaag caccgtgaaa ccctggaccc	780
cagcgcccc aaggacctca tcgacaccta cctgctccac atggaaaaag agaaatccaa	840
cgcacacagt gaattcagcc accagaacct caacctcaac acgctctcgc tcttctttgc	900
tggcactgag accaccagca ccactctccg ctacggcttc ctgctcatgc tcaaataccc	960
tcatgttgca gagagagtct acagggagat tgaacagggt attggccac atcgccctcc	1020
agagcttcat gaccgagcca aaatgccata cacagaggca gtcattctatg agattcagag	1080
attttccgac cttctcccca tgggtgtgcc ccacattgtc acccaacaca ccagcttccg	1140
agggtagatc atccccaagg acacagaagt atttctcatc ctgagcactg ctctccatga	1200
cccacactac tttgaaaaac cagacgcctt caatcctgac cactttctgg atgccaatgg	1260
ggcactgaaa aagactgaag cttttatccc cttctcctta ggggaagcga tttgtcttgg	1320

tgaaggcatc	gcccgtgcgg	aattgttcct	cttcttcacc	accatcctcc	agaacttctc	1380
catggccagc	cccgtggccc	cagaagacat	cgatctgaca	ccccaggagt	gtggtgtggg	1440
caaaatacc	ccaacatacc	agatccgctt	cctgccccgc	tgaaggggct	gaggggaaggg	1500
ggtcaaagga	ttccagggtc	attcagtgtc	ccgcctctg	tagacaatgg	ctctgactcc	1560
ccgcaacttc	ctgcctctga	gagacctgct	acaagccagc	ttccttcccc	tccatggcac	1620
cagttgtctg	aggtcacatt	gcaagtgagt	gcaggagtga	gattatcgaa	aattataata	1680
tacaaaatca	tatatatata	tatgttcttg	ttttttgaga	cagagtctca	cactgttgcc	1740
caggctggag	tgcagtggcg	tgatctcggc	tactgcaac	ctccaccccc	ggggatcaag	1800
caactctcct	gcctcagcct	ccctagtagc	tgggattaca	ggcatgcact	accacgcttg	1860
gctaattttt	gtatttttag	tagagatggg	gtttcactgt	gtaggccagg	ctggtctcga	1920
actcctgaac	tcaagtgatt	caccacactt	agcctcccaa	agtgtctggga	ttacaggcgt	1980
gagtcaccgt	gcccagccat	gtatatatat	aatttttaaaa	attaagctga	aattcacata	2040
acataaaatt	agctgtttta	aagtgtaaaa	tttagtggcg	tgtggttcat	tcacaaagct	2100
gtacaaccac	caccatctag	ttccaaacat	tttctttttt	tctgagatgg	agtctcactc	2160
tgtcacccag	gttcgagttc	agtggtgcca	tctctgtcca	ctgcaacctc	cacatcctgg	2220
gttcaagtga	ttctcctgcc	tcagcctctg	gaggagctgg	tatcacaggc	gtccccccacc	2280
acgcctggct	aaatttttga	tttttaggtg	gtcttgaact	cctgatgtca	ggtgattctc	2340
ctagctccaa	atgtttttcat	tatctctccc	ccaacaaaac	ccatacctat	caagctgtca	2400
ctccccatac	cccattctct	ttttcatctc	ggccccctgtc	aatctggttt	ttgtcactat	2460
ggacttacca	attctgaata	tttcccataa	acagaatcat	acaatatattg	attttttttt	2520
tttttttgaa	actaagcctt	gctctgtctc	ccaggctgga	gtgctatggg	gcaatttttg	2580
ttcactgcaa	cctctgcctt	ccaagatcaa	gagattctcc	agtctcagct	cccaagtagc	2640
tgggattaca	ggcatgtact	accatgcctg	gctaattttc	ttgtagtttt	agtagggaca	2700
tgttggccag	gctgggtggg	agctcctggc	ctcagggtgat	ccacccacct	cagtgttcca	2760
aagtgtgat	attacaggca	taatatgtga	tcttttgtgt	ctggttgctt	tcatgttgaa	2820
tgctattttt	gaggttcatt	cctgtttag	accacagtca	cacactgctg	tagtcttccc	2880
cagtcctcat	tcccagctgc	ctcttcctac	tgcttccgtc	tatcaaaaag	cccccttggc	2940
ccaggttccc	tgagctgtgg	gattctgcac	tggtgctttg	gattccctga	tatgttcctt	3000
caaactctgct	gagaattaaa	taaacatctc	taaagcctga	cctccccacg	tc	3052

<210> 47

<211> 1645

<212> DNA

<213> NM\_003811.2| Homo sapiens tumor necrosis factor (ligand) superfamily, member 9 (TNFSF9), mRNA

```

<400> 47
agtctctcgt catggaatac gcctctgacg cttcactgga ccccgagcc cctgggcctc 60
ccgcgccccg cgctcgcgcc tgccgcgtac tgccttgggc cctggtcgcg gggctgctgc 120
tgctgctgct gctcgcctgcc gcctgcgcgc tcttcctcgc ctgcccctgg gccgtgtccg 180
gggctcgcgc ctgcgccggc tccgcggcca gcccagagact ccgcgagggg cccgagcttt 240
cgccccagca tcccgcggc ctcttgacc tgcggcaggg catgtttgcg cagctggtgg 300
cccaaatgt tctgctgac gatgggcccc tgagctggta cagtgaacca ggcctggcag 360
gcgtgtccct gacggggggc ctgagctaca aagaggacac gaaggagctg gtggtggcca 420
aggctggagt ctactatgtc ttctttcaac tagagctgcg gcgcgtgggt gccggcgagg 480
gctcaggctc cgtttcactt gcgctgcacc tgcagccact gcgctctgct gctggggccg 540
ccgccttggc ttgaccgtg gacctgccac ccgcctcctc cgaggctcgg aactcggcct 600
tcggtttcca gggccgcttg ctgcacctga gtgccggcca gcgcctgggc gtccatcttc 660
aactgaggc cagggcacgc catgcctggc agcttaccca gggcgccaca gtcttgggac 720
tcttcgggt gacccccgaa atcccagccg gactcccttc accgaggtcg gaataacgcc 780
cagcctgggt gcagcccacc tggacagagt ccgaatccta ctccatcctt catggagacc 840
cctggtgctg ggtccctgct gctttctcta cctcaagggg cttggcaggg gtccctgctg 900
ctgacctccc cttgaggacc ctccctaccc actccttccc caagttggac cttgatattt 960
attctgagcc tgagctcaga taatatatta tatatattat atatatatat atatttctat 1020
ttaaagagga tcctgagttt gtgaatggac ttttttagag gagttgtttt gggggggggg 1080
tcttcgacat tgccgaggct ggtcttgaac tcctggactt agacgatcct cctgcctcag 1140
cctcccaagc aactgggatt catcctttct attaatcat tgtacttatt tgcctatttg 1200
tgtgtattga gcatctgtaa tgtgccagca ttgtgccag gctagggggc tatagaaaca 1260
tctagaaata gactgaaaga aaatctgagt tatggtaata cgtgaggaat ttaaagactc 1320
atccccagcc tccacctcct gtgtgatact tgggggctag cttttttctt tctttctttt 1380
ttttgagatg gtcttgttct gtcaaccagg ctagaatgca gcggtgcaat catgagtcaa 1440
tgcagcctcc agcctcgacc tcccagggt caggatgatcc tcccatctca gcctctcgag 1500
tagctgggac cacagttgtg tgccaccaca cttggctaac tttttaattt ttttgcgag 1560
acggtattgc tatgttgcca aggttgttta catgccagta caatttataa taaacactca 1620
tttttcctca aaaaaaaaaa aaaaa 1645

```

<210> 48

<211> 6640

<212> DNA

<213> NM\_006047.4| Homo sapiens RNA binding motif protein 12 (RBM12), transcript variant 1, mRNA

<400> 48  
gagcgggccc acggccattt tgtgaagcgg cgaaggaggt ggtggctgcg ttgggctccg 60  
ggaagccgtt cgggctgggg ctgtcggccg cggggcggag gcactcgcgc ggggggtaat 120  
tcgggggtctg ggttctgggt ccgcgcagct ttccccgtct aaaagttggt ttttaattggt 180  
tgcccacagg attgacttgg cctctacttc ttgttaagga aattcatctc ttgtttttatc 240  
agggtgtgtgt ggtttcagcg cagcatggct gtggctcatcc gtttgcaagg tctcccaatt 300  
gtggcgggga ccatggacat tcgccacttc ttctctggat tgaccattcc tgatggggggc 360  
gtgcatattg taggggggtga actgggtgag gctttcatcg tttttgccac tgatgaagat 420  
gcaaggcttg gtatgatgcg cacagggtgg acaattaaag ggtcaaaagt aacactattg 480  
ttgagtagta agacggaaat gcagaatatg attgaactga gtcgtaggcg ttttgaaact 540  
gccaacttag atataccacc agcaaatacc agtagatcag gaccaccacc tagctcagga 600  
atgagtagca gggtaaactt gccacaaca gtatccaact ttaataatcc atcaccagat 660  
gtagttactg ccaccacttc tgttcatgaa agcaacaaaa acatacagac attttccaca 720  
gccagcgtag gaacagctcc tccaaatatg ggggcttcc tggggagccc aacgtttagc 780  
tcaactgttc caagcacagc ctctccaatg aacacagtc cgcgcgccacc aattcctcca 840  
attccagcga tgccatctct gccaccaatg ccatccattc cccaattcc agttcctcct 900  
ccagtaccta cattgcctcc tgtgcctcct gtgccccga tccccccagt tccttctgtg 960  
ccaccatga cccactgcc acccatgtcg ggcagtcgc ccttgaatcc gccacctgtg 1020  
gcacctctac ctgctggaat gaatggctct ggagcaccta tgaatttgaa caataatctg 1080  
aatcctatgt ttcttgggtc gttgaatcct gttaacccta tccagatgaa ctctcagagc 1140  
agtgtgaagc cactcccat caaccctgat gatctgtatg tcagtgtgca tggaatgccc 1200  
ttttctgcaa tggaaaatga tgtcagagat ttttttcatg ggctccgtgt tgatgcagtg 1260  
catttgttga aagatcatgt aggtcgaaat aatgggaatg gattgggttaa gtttctctcc 1320  
cctcaagata catttgaaagc tttgaaacga aacagaatgc tgatgattca acgctatgtg 1380  
gaagttagcc ctgccacaga aagacagtgg gtagctgctg gaggccatat cacttttaag 1440  
caaaatatgg gaccttctgg acaaactcat cccctcctc agacacttcc cagggtcaaaa 1500  
tcgcccagtg ggcagaaaag atcaaggta agatcaccac atgaggctgg tttttgtgtt 1560  
tacttgaaag ggctaccatt tgaagcagaa aacaaacatg tcattgattt ttttaaaaag 1620  
ctggatattg tggaagatag tatttatata gcttatggac ccaatgggaa agcaactggc 1680  
gaaggctttg tagagttcag aaatgaggct gactataagg ctgctctgtg tcgtcataaa 1740  
cagtacatgg gcaatcgctt tattcaagtt catccaatta ctaagaaagg tatgctagaa 1800

aagatagata	tgattcgaaa	aagactgcag	aacttcagct	atgaccagag	ggaaatgata	1860
ctaaatccag	agggggatgt	caactctgcc	aaagtctgtg	cccacataac	aaatattcca	1920
ttcagcatta	caaagatgga	tgttcttcag	ttcctagaag	gaatcccagt	ggatgaaaat	1980
gctgtacatg	ttcttgttga	taacaatggg	caaggctctag	gacaggcatt	ggttcagttt	2040
aaaaatgaag	atgatgcacg	taagtctgaa	cgcttacacc	gtaaaaaact	taatgggaga	2100
gaagctttttg	ttcatgtagt	taccctagaa	gatatgagag	agattgagaa	aaatccccct	2160
gccaagggaa	aaaagggatt	aaagatgcct	gtgccaggta	atcctgcagt	tccaggaatg	2220
cccaatgcgg	gactgcccgg	tgtgggactg	cccagtgcag	gacttcccgg	tgcaggcctg	2280
cccagcacag	gactgcctgg	ttcagcaata	accagtgcag	gactgcctgg	tgcgggaatg	2340
cccagtgcag	gaatacctag	tgcaggaggt	gaagagcatg	ccttcctgac	tgtaggatca	2400
aaggaagcca	ataatggggc	tccatttaac	tttcttggtg	attttggtgg	atcaaagcc	2460
tttggggccac	caatccctcc	tccaggatta	ggaggcgggg	cctttggtga	tgctaggcct	2520
ggtatgcctt	cagttggaaa	cagtggtttg	cctggtctag	gactggatgt	tccgggtttt	2580
ggaggtggac	caaacaattt	aagtgggcca	tcgggatttg	gagggggccc	tcagaatttt	2640
ggaaatggcc	ctggtagctt	aggcgggtccc	ccgggggttg	gaagtggccc	tcctggtcct	2700
ggaagtggcc	ctgggcattt	gggtgggcca	ccagcttttg	ggcctggccc	cggccccggc	2760
cccggccctg	gccaatcca	tattggtggt	ccccctggct	ttgcatctag	ttctggaaaa	2820
ccaggaccga	cagtaattaa	agtgcaaaac	atgcccttta	ctgtgtctat	tgatgagatt	2880
ttagatttct	tttatggcta	tcaagtaatc	ccaggctcag	tgtgtttaaa	atacaatgaa	2940
aaaggtatgc	ccacagggtga	agccatggtg	gcctttgagt	ctcgggatga	agccacagct	3000
gctgtcattg	acttaaatga	caggcctata	ggttcaagaa	aagtaaaact	tgtattaggg	3060
tagccattca	catcatTTTT	tatagggtag	atcttcatat	tgctgtgatt	aatgcatcca	3120
gattgttttc	ctagtatttc	caggttagaa	cctgtggatt	gtttcaattg	catatagctt	3180
ggtttccata	acatagagca	ttggttgact	gtttacagaa	gactcactca	ccaggataaa	3240
cattgctgta	tgttacagta	aagctatctg	gagagaacac	ataaatgatt	ttggcatacc	3300
attagagaaa	ccatttgtaa	aactcaaatg	accacataaa	gcttatcaag	gagtctagat	3360
tggttttggt	ttataccata	tgggatgaag	aaaatagaaa	tgtcagtaga	actcattgag	3420
ggtgctcttg	ccagctgctg	aaaatagaag	ttggctactc	tcagaatttg	gtttaaagct	3480
ggacagattt	gctttgttat	agggtaaagc	tttgtctaaa	gtcctcattt	tcttttaaaa	3540
ttgaataaaa	tttctgtata	cagattcatt	gtatgtacct	ttattgcttc	ttaagggctc	3600
ttgctgtata	gacagtcctg	cttcagaagt	tgctgctttg	tttgtcta	tgactcattt	3660
gtaaatgagc	agaactgttt	tgttggtttt	tttccctaaa	tataaaagtc	cacacttcgt	3720
ttgtgctata	acctcaaaact	ttgattttcta	atgtcacact	taaaactgtg	tggaataaga	3780

cttttgccat	aaaaataaac	tatggagtc	tttatctacc	agagcctttt	tggtttgacc	3840
gccacgattt	aggttagtca	gtttaaaaat	tgttcatgtt	gtttggatgg	tatcgaaaac	3900
cagaaaccac	ttttaataat	gtgtttaaga	tacttgattt	gaagtccttt	tcatatggac	3960
taattgtagt	atcaattttc	tcctgtcccg	attatgtgaa	attttggcct	ttaaacaaag	4020
agggggcccat	tcataagaaa	gtgttatatc	taggttttta	aaactgaagt	tgaaattatc	4080
tttgtttagca	gtagtagtat	agaataaaaag	atccgtatgc	tggttcgtag	attgatacgt	4140
gttagtcctg	ttatttggag	gctttttggc	atagttgttc	gatcaggagc	ctgtttacta	4200
aaagtcttca	tacagagtac	aagtgcagcc	gccagaggag	aaaattgaga	ttcttgaccc	4260
tttcatactc	ttttctttgg	tattcaggac	actaaggcag	gaggaccaca	tgagttctgg	4320
ttggtaagtg	tccttgtcat	gaaaacactt	gccttacaag	gctctaatta	tgatttcctt	4380
agtcagtgct	ctgaagatgt	gtcacattat	attataaaca	atacggaagg	ggagataggt	4440
gagatgatat	gaaaaacaaa	ttttctcact	gtcataaaag	gctctaatta	tggttacttt	4500
ccttgtgatg	aaaaacttgg	ccttacaggg	ctctagttat	ggttactttc	cctagtcaat	4560
gctctgaaaa	tgtgtcacat	tataaacaat	acagaaggag	agatagatga	gagatcatga	4620
aaaccaaatt	ttatctttta	catggccctt	ttgtcttcgt	ttgaaacatc	cccaacattt	4680
cctacaaatc	agcgtagtta	caaaggggtc	agtcttttaa	aataagtatt	tcctattaaa	4740
ctatatatat	atacagtgcc	tttttggtgt	tgtgagtcag	tggaacactg	aaatacagcg	4800
gttgtgtaat	ttaagagtgg	caacagtttc	atttgataag	atttgaaaag	gctttttatc	4860
actacaatct	tagaggattg	acagtacagg	atttttgttt	aagagaagga	ttgttttagac	4920
tcaagaggtg	actatgttgt	gggtcttttg	ataattcatg	aatacagatt	tgctttgacc	4980
gatcactaga	tactgcctcc	tcaatttcaa	aagcaatata	acgtttgtat	atgctgttta	5040
atttaagtta	atgttaagta	atcattttct	attccaaaga	caatgcaaaa	aacttcagca	5100
ctgcttgaga	gttgatattt	agccacagat	attttctcca	ttggaaagct	attttcattt	5160
tagaaaaatg	ggttgtttga	agatgaaagt	cttttatctt	ttttcacaat	ttattttggg	5220
tatatgttcg	tacattctta	ttaaataatt	catacacttt	attgcaacta	ctttgttatt	5280
tctacatctt	agataaatgc	tgaaaaggaa	aacgatttca	ttgttcatct	aattaaataa	5340
attaaaagag	cggtttgtgt	agaaattgga	gagaactatg	ttttatatga	atcacaacag	5400
tgctcgttgt	ggactgtaat	taaatgcttt	gccctctgga	acttgatttt	tgtgtatctg	5460
agatttatta	acagtgtctaa	ctgctaagga	tactgtttat	cttgttctgg	gcattggagt	5520
ttcagttttt	aaaaaattct	tgaaaatgta	ttctgtgaaa	ctactcatac	ctctcttcct	5580
gcgaattttc	tcctaagaac	taggtttggg	gtagaaatga	attgacatta	ttttctcatt	5640
tgctttgtat	ggatgaagg	atttgtaaag	ctttgtctca	agttgtgtgc	atttgtaaac	5700
actatcatga	tattttcaat	tttatgtgta	aattttattg	tctgtttttg	gtgactctga	5760
cattaatgga	agagaatatt	ttccattaga	tttaattttt	tttcctctcc	cttctgattt	5820

tttttttatt ggtgttcatt tttcttttga tttaaaggat tagagaaatc tacaaatgta	5880
tgtcataaat aagcaaattt gtaaactttt ctgaacttta gtgaaacatt tagttcacaa	5940
gcataatttg gaggtgtggt ctgtgtttac acagtagttt ggatgtacaa ttattattag	6000
tggcttttta aaaaatgaaa cagtgttaag tgaaatgtag ttcctagctt tgtactccaa	6060
gttgtcaaag catcaacaat gaaaattcga ttaggaaact ttatttataaa tttcaggtag	6120
taatattcag tgtagttaag gccagtctta acccactgga tgaaaatcta ggactgtatg	6180
gaagtaagca aacattacat ttttaggttg aaatagtcag ccttgcataa aaacaaggat	6240
gcgtgaaagc cttaaattcc agctcccttt tactggagtc tgtggttggtg tacaggtag	6300
ggccaagtgt aaaatctcat caattttaag aacactcggg aaatgagtaa agaaaatgta	6360
aaattgctgc tagtcaaatc ttttggaag aatttctgga agtgggtcact ttaaaaatta	6420
ttttccaccc ttgcaaaatt gccacattta aattgtttta ctggcagttc tatagtagtc	6480
cagacttttag aaaccaaaca caacaaaatg gcttggtgcc aatatggcca caacattgcc	6540
agcaaatact gccttggcat cactcagcag aggtttttgt ttataaagat gaagtcttga	6600
atactgttca ataaacttgt caaaaaataa aaaaaaaaaa	6640

<210> 49

<211> 3680

<212> DNA

<213> NM\_006644.2| Homo sapiens heat shock 105kDa/110kDa protein 1 (HSPH1), mRNA

<400> 49	
tgagtaaata cgcgagattc tggaaagttc tgatcagtc gatacataag gctgaggaag	60
tgggacctcc ctttttgggt cggtagttca gcgccggcgc cgggtgtgca gccgcggcag	120
agtgaggcag gcaacccgag gtgcggagcg acctgcggag gctgagcccc gctttctccc	180
agggtttctt atcagccagc cgccgctgtc cccgggggag taggaggctc ctgacaggcc	240
gcggctgtct gtgtgtcctt ctgagtgta gaggaacggc cagaccccgc gggccggagc	300
agaacgcggc cagggcagaa agcggcggca ggagaagcag gcagggggcc ggaggacgca	360
gaccgagacc cgaggcggag gcggaccgcg agccggccat gtcggtggtg gggttggacg	420
tgggctcgca gagctgctac atcgcggtag cccggggccg gggcatcgag accatcgcca	480
atgagttcag cgaccggtgc accccgtcag tcatatcatt tggatcaaaa aatagaacaa	540
tcggagttgc agccaaaaat cagcaaatca ctcatgcaaa caatacgggtg tctaacttca	600
aaagatttca tggccgagca ttcaatgacc ctttcattca aaaggagaag gaaaacttga	660
gttacgattt ggttccattg aaaaatggtg gagttggaat aaaggtaatg tacatgggtg	720
aagaacatct atttagtgtg gagcagataa cagccatgtt gttgactaag ctgaaggaaa	780



ctgctgaaaa	cagcctcaag	aaaccagtaa	cagattgtgt	tatttcagtc	ccctccttct	840
ttacagatgc	tgagaggcga	tctgtgttag	atgctgcaca	gattgttggc	ctaaactggt	900
taagacttat	gaatgacatg	acagctgttg	ctttgaatta	cggaatttat	aagcaggatc	960
tcccaagcct	ggatgagaaa	cctcggatag	tggtttttgt	tgatatggga	cattcagctt	1020
ttcaagtgtc	tgcttgtgct	tttaacaagg	gaaaattgaa	ggtactggga	acagcttttg	1080
atcctttctt	aggaggaaaa	aacttcgatg	aaaagttagt	ggaacatttt	tgtgcagaat	1140
ttaaaactaa	gtacaagttg	gatgcaaaat	ccaaaatacg	agcactccta	cgtctgtatc	1200
aggaatgtga	aaaactgaaa	aagctaata	gctctaacag	cacagacctt	ccactgaata	1260
tcgaatgctt	tatgaatgat	aaagatgttt	ccggaaagat	gaacagggtca	caatttgaag	1320
aactctgtgc	tgaacttctg	caaaagatag	aagtaccctt	ttattcactg	ttggaacaaa	1380
ctcatctcaa	agtagaagat	gtgagtgcag	ttgagattgt	tggaggcgct	acacgaattc	1440
cagctgtgaa	ggaaagaatt	gccaaattct	ttggaaaaga	tattagcaca	acactcaatg	1500
cagatgaagc	agtagccaga	ggatgtgcat	tacagtgtgc	aatactttcc	ccggcattta	1560
aagttagaga	atcttccgtc	acagatgcag	ttccttttcc	aatatctctg	atctggaacc	1620
atgattcaga	agatactgaa	ggtgttcatt	aagtccttag	tcgaaacat	gctgctcctt	1680
tctccaaagt	tctcaccttt	ctgagaaggg	ggccttttga	gctagaagct	ttctattctg	1740
atccccaagg	agttccatat	ccagaagcaa	aaataggccg	ctttgtagtt	cagaatgttt	1800
ctgcacagaa	agatggagaa	aaatctagag	taaaagtcaa	agtgcgagtc	aacacccatg	1860
gcattttcac	catctctacg	gcatctatgg	tggagaaagt	cccaactgag	gagaatgaaa	1920
tgtcttctga	agctgacatg	gagtgtctga	atcagagacc	accagaaaac	ccagacactg	1980
ataaaaaatgt	ccagcaagac	aacagtgaag	ctggaacaca	gccccaggta	caaactgatg	2040
ctcaacaaac	ctcacagtct	cccccttcac	ctgaacttac	ctcagaagaa	aacaaaaatcc	2100
cagatgctga	caaagcaaat	gaaaaaaaag	ttgaccagcc	tccagaagct	aaaaagccca	2160
aaataaaggt	ggtgaatgtt	gagctgccta	ttgaagccaa	cttgggtctgg	cagttaggga	2220
aagaccttct	taacatgtat	attgagacag	agggttaagat	gataatgcaa	gataaattgg	2280
aaaaagaaaag	gaatgatgct	aaaaatgcag	ttgaggaata	tgtgtatgag	ttcagagaca	2340
agctgtgtgg	accatatgaa	aaatttatat	gtgagcagga	tcacaaaaat	tttttgagac	2400
tcctcacaga	aactgaagac	tggctgtatg	aagaaggaga	ggaccaagct	aaacaagcat	2460
atgttgacaa	gttggaaagaa	ttaatgaaaa	ttggcactcc	agttaaagtt	cggtttcagg	2520
aagctgaaga	acggccaaaa	atgtttgaag	aactaggaca	gaggctgcag	cattatgccca	2580
agatagcagc	tgacttcaga	aataaggatg	agaaatacaa	ccatattgat	gagtctgaaa	2640
tgaaaaaagt	ggagaagtct	gttaatgaag	tgatggaatg	gatgaataat	gtcatgaatg	2700
ctcaggctaa	aaagagtctt	gatcaggatc	cagttgtacg	tgctcaggaa	attaaaacaa	2760

aaatcaagga attgaacaac acatgtgaac ccgttgtaac acaaccgaaa ccaaaaattg	2820
aatcacccaa actggaaaga actccaaatg gcccaaatat tgataaaaag gaagaagatt	2880
tagaagacaa aaacaatttt ggtgctgaac ctccacatca gaatggtgaa tgttacccta	2940
atgagaaaaa ttctgttaat atggacttgg actagataac cttaaattgg cctattcctt	3000
caattaataa aatatTTTTTg ccatagtatg tgactctaca taacatactg aaactattta	3060
tattttcttt ttttaaggata tttagaaatt ttgtgtatta tatggaaaaa gaaaaaaagc	3120
tttaagtctgt agtcttttatg atcctaaaag ggaaaattgc cttggtaact ttcagattcc	3180
tgtggaattg tgaattcata ctaagctttc tgtgcagtct caccatttgc atcactgagg	3240
atgaaactga cttttgtctt ttggagaaaa aaaactgtac tgcttgttca agagggctgt	3300
gattaaaatc ttttaagcatt tgttcctgcc aaggtagttt tcttgcatth tgctctccat	3360
tcagcatgtg tgtgggtgtg gatgtttata aacaagacta agtctgactt cataagggct	3420
ttctaaaacc atttctgtcc aagagaaaat gactttttgc tttgatatta aaaattcaat	3480
gagtaaaaca aaagctagtc aaatgtgtta gcagcatgca gaacaaaaac tttaaacttt	3540
ctctctcact atacagtata ttgtcatgtg aaagtgtgga atggaagaaa tgtcgatcct	3600
gttgtaactg attgtgaaca cttttatgag ctttaaaata aagttcatct tatggtgtca	3660
tttctaaaaa aaaaaaaaaa	3680

<210> 50

<211> 3349

<212> DNA

<213> NM\_004602.1| Homo sapiens staufen, RNA binding protein (Drosophila) (STAU), transcript variant T4, mRNA

<400> 50	
acttcctgcc gggctgcggg cgcctgagcg ctcttcagcg tttgcgcggc ggctgcgcgt	60
ctctctcggc tcccgcttcc tttgaccgcc tccccccccc ggccccggcg cgccccctc	120
ctccacggcc actccgcctc ttccctccct tcgtcccttc ttctctccc tttttcctt	180
cttcttccc ctctcgcgcg ccaccgcccc ggaccgcccg ccgggggacg agtccggagc	240
agcagccagc agcagccagg tggagttttg ctcttgctgc ccaggctgga gtgcagtggc	300
gtgatctcgg ctactgcaa cctccacctc ccagggtcagc gattttccca cttcagcctc	360
ccgataagct gagattacag gagtttatta accacttaac ctctcagaac tgaacaaaga	420
caacattggt cctggaacgc cctcttttta aaaaagaaag cataaccctt actgtagaac	480
taaattgcact gtgcatgaaa cttggaaaaa aaccaatgta taagcctgtt gacccttact	540
ctcggatgca gtccacctat aactacaaca tgagaggagg tgcttatccc ccgagggtact	600
tttaccatt tccagttcca cttttacttt atcaagtgga actttctgtg ggaggacagc	660

aattttaatgg	caaaggaaaag	acaagacagg	ctgcgaaaca	cgatgctgct	gccaaaagcgt	720
tgaggatcct	gcagaatgag	cccctgccag	agaggctgga	ggtgaatgga	agagaatccg	780
aagaagaaaa	tctcaataaa	tctgaaataa	gtcaagtgtt	tgagattgca	cttaaacgga	840
acttgccctgt	gaatttcgag	gtggccccggg	agagtggccc	acccacatg	aagaactttg	900
tgaccaaggt	ttcggttggg	gagtttgttg	gggaagggtga	agggaaaagc	aagaagattt	960
caaagaaaaa	tgccgccata	gctgttcttg	aggagctgaa	gaagttaccg	cccctgcctg	1020
cagttgaacg	agtaaagcct	agaatcaaaa	agaaaacaaa	acccatagtc	aagccacaga	1080
caagcccaga	atatggccag	gggatcaatc	cgattagccg	actggcccag	atccagcagg	1140
caaaaaagga	gaaggagcca	gagtacacgc	tcctcacaga	gcgaggcctc	ccgcgccgca	1200
gggagtttgt	gatgcagggtg	aaggttggaa	accacactgc	agaagggaacg	ggcaccaaca	1260
agaagggtggc	caagcgcaat	gcagccgaga	acatgctgga	gatccttggt	ttcaaagtcc	1320
cgcagcggca	gcccaccaa	cccgcactca	agtcagagga	gaagacaccc	ataaagaaac	1380
caggggatgg	aagaaaagta	accttttttg	aacctggctc	tggggatgaa	aatgggacta	1440
gtaataaaga	ggatgagttc	aggatgcctt	atctaagtca	tcagcagctg	cctgctggaa	1500
ttcttcccat	ggtgcccag	gtcgcgccag	ctgtaggagt	tagtcaagga	catcacacca	1560
aagattttac	cagggcagct	ccgaatcctg	ccaaggccac	ggtaactgcc	atgatagccc	1620
gagagttggt	gtatgggggc	acctcgccca	cagccgagac	cattttaaag	aataacatct	1680
cttcaggcca	cgtaccccat	ggacctctca	cgagaccctc	tgagcaactg	gactatcttt	1740
ccagagtcca	gggattccag	gttgaataca	aagacttccc	caaaaacaac	aagaacgaat	1800
ttgtatctct	tatcaattgc	tcctctcagc	cacctctgat	cagccatggt	atcggcaagg	1860
atgtggagtc	ctgccatgat	atggctgcgc	tgaacatctt	aaagttgctg	tctgagttgg	1920
accaacaaag	tacagagatg	ccaagaacag	gaaacggacc	aatgtctgtg	tgtgggaggt	1980
gctgaacctt	ttctggccat	gaaccattat	aaaatcccaa	catatatact	gaaaatactg	2040
aaactgcttt	gaaaatttgg	aatttctgat	acctccagt	ggccgagaga	cacggtgggt	2100
aaaggatgtg	ggcagcagca	gggaagacaa	cagaaacaca	aggaggcggc	tgtggccggc	2160
tggactgtgc	tggggtttgt	tgtgatggcc	actcggtgac	ctggcgggtcc	ctacgcaata	2220
gcagctgcct	gtggggaaga	agggtgccc	agccagctgg	ttctcccggg	acaccagcag	2280
atccacaccc	tgggcacctc	cgtgtttggt	cttttttttc	ccctgtgtga	aagaagaaac	2340
ggcacgaccc	cttctcaagc	tggtctactc	agacacattg	ggacaaaccc	tggacagcca	2400
tgccagagag	aggcctttga	ccggccccag	agctaaaagc	accagagaaa	atcaaattgct	2460
tcctactcag	cgtgacccaa	cttttctagt	gtgccacggc	cccaccacct	cctgcagtac	2520
ccacaccatc	accactgctt	tctcttccaa	cagtgatctg	tattcttagt	ttcattattt	2580
tcttttgatt	gatatgacac	tatataaaat	tttcatttga	gaatttctca	attgtatcta	2640
gttaaatagc	acagtttggg	aacttgtctg	agactgactt	tatcaataat	ctaaccgaca	2700

aagatcatat ccatgtgtat gtggtagac atttttat	2760
tttcagtgat gcaaattgtg tgccctctgg ttcagctgaa	2820
accttgaata agtctccac agttgtataa attggacaat	2880
atgatcattt ggttccattt ttatttcatt tttat	2940
atgaactttg atctctgttt taaagattat taaaaaacat	3000
tcttgaggac ttagctttca ctacactaca ggatatgac	3060
ctgcagagtg attttccaga gtgctcgata ctgttaatta	3120
agaatgacct acgtttctgt atacagctgt gttgctttt	3180
gaagtgtgtg cactgaggct ctgctgtgtg tccgtatgga	3240
ttaagtactg cttccattca ttgtttacgc tggaatttt	3300
aaacttaagt gtttgtcatc aataaatggt aatactaaaa	3349

<210> 51

<211> 402

<212> DNA

<213> NM\_021246.2| Homo sapiens lymphocyte antigen 6 complex, locus G6D (LY6G6D), transcript variant 1, mRNA

<400> 51	
atgaaacccc agtttgttgg gatcttgctc agctccctgc	60
cgaatgcggt gctacaactg tgggtggaagc cccagcagtt	120
acctgtggcg agggcagacc ccagccaggc ctggaacaga	180
ccagtgacct tgattcacca acatccagcc tgcgtcgag	240
gagacagagt cggtgggaga cgtgacttat ccagcccaca	300
ctgtgcaaca gcgccgtggc aagccatgtg gccctgcag	360
accgccctga cctgtctctt gccaggactg tggagcggat	402

<210> 52

<211> 3248

<212> DNA

<213> NM\_007236.3| Homo sapiens calcium binding protein P22 (CHP), mRNA

<400> 52	
accacccctg gggtccctcc cgggtccgca gtggaacac	60
cctagccctt cttccctcc ctcttccct cctgtcgccg	120
cccggaggag ctcccggcac ggcgatgggt tctcgggcct	180

gagctcgagg agatcaagaa ggagaccggc ttttcccaca gtcaaatacac tgcctctac	240
agccggttca ccagcctgga caaaggagag aatgggactc tcagccggga agatttccag	300
aggattccag aacttgccat caaccactg ggggaccgga tcatcaatgc cttctttcca	360
gagggagagg accaggtaaa cttccgtgga ttcattgcga ctttggtca tttccgcccc	420
attgaggata atgaaaagag caaagatgtg aatggaccg aaccactcaa cagccgaagc	480
aacaaactgc actttgcttt tcgactatat gatttgata aagatgaaa gatctcccgt	540
gatgagctgt tacagggtgt acgcatgatg gtcggagtaa atatctcaga tgagcagctg	600
ggcagcatcg cagacaggac cattcaggag gctgatcagg atggggacag tgccatatct	660
ttcacagaat ttgttaaggt tttggagaag gtggatgtag aacagaaaat gagcatccga	720
tttcttcaact aaaggagacc aaactgttcc ttgcggtcta gtatttaaga actggaactt	780
gaaagtcctc cttctacca ctccacctcc acccctcat tccccttctc ccaaagtact	840
actgctgttg catgacaacc ccaaataatgt tctgtcaaca caaacctgcc tttggtgtat	900
aaacagggca ttacagaatg gtacacctta tatatttctg ttcagtatcc attcactagt	960
tcttcattta taaatatcat cttccccatt ctgctgtga atgccacaca tccatccagt	1020
ctgagaaagt gagagaggca atcatgccaa gaacaagcca gcaaagctct ttcaccagat	1080
gtagactgta gccctgctgc cttccctcca gcgagtctgc cagcatgctt cttcatcctt	1140
tttatatgtt ctttgcttcc tacttccctg tcttccaaca tactgttcac ttactctggc	1200
agtctttctg cttttcatta agcctcaaaa tctcctctgt tctacttggc accacaagct	1260
atgtcctata tatgtatttc tgacttggca ggatagttca ggggtctggc agtttttatt	1320
taccttcatt attaaatggg cctctgggat gttgcctctt caggagcttt ttggtaatca	1380
atacttctct cagaagtatg agaccatcct ctgcactctg ctctgtcatc aaaggctgct	1440
gggtggagat accctttttg aaagggtggc ttggtgagag gtatggagcc aagtcttcta	1500
ggttgcttgc ccacatcact ctatctctgg cctctgattc tcaactttgt acctgtgtgg	1560
ctcctcttgt tagtgcaatg ttgactgttg aaaaagcagc agtatgctta caggtttgc	1620
tagtttggg acaccgttac caccagaatg gctgctctga caatatgcct agggactttc	1680
tcatggcttt ttttaataa ggaggctggg caccctataa agcctcatgc attcacacct	1740
ttgcagcatg gtttatgcct cagtgttatg tgcactggaa tgttttccac ttcacatttc	1800
caagtagaaa tattagtgtt acggaagtgc ctaatatccc agtccaaatt tttttttttt	1860
tttttttttt tttttgagac agagtcttgc tctgtcacc aggctggagt gcagtgggtgc	1920
gatcgctcac tgcaacctca gcctcctgga ttttaagtgt tctcctgcct cagcctccca	1980
agtagctggg attacagggtg tgcaccacca tgcccggcta attttttgta tttttagtgg	2040
agacagggtt tcaccatgtt ggccaggctg gtctcgaact cctgacctcg tgatccgcct	2100
gcctcagcct cccaaagtgc tgggattaca ggtgtgagcc accacgcctg gccccagtcc	2160

```

aaaatattta aagattgttt ccttagtgtc ttgaagtttt gcacaaaatt cttttttttg 2220
agatggagtc tcaactctgtc acccaggctg gagtgcagtg gcgtgatctt ggctcactgc 2280
aacctctgcc tcctgggttc aagcaattct cccacctcag cctcccaagt agctgggatt 2340
acagacgtgt gccaccatac ctgggtaatt tttgcatttt tagtggagag ggagtttcac 2400
catgttggcc aggttggctt tgaactcctg acctcaggtg atcctcctgc ctcggcctcc 2460
caaagtgtg ggattacagg catgagccac cgtgctcagc cgcaaaattc tttatgaatt 2520
ttacacttgg caaatgttaa tgacggaagc catagtctgc tcctaataca tgtccaaagc 2580
attgactgtt gtgtcattag ctgcctggtt acattagctc cctggcttct tgttttagacc 2640
actgctaadc ctttaaaaac aagaggtctg gcactagtag cacaacctaa ggtggcatta 2700
cagatctttg agcgagccac agcaactttt ctgccaagtc agcttagttt agacttcagt 2760
gaatcaggct attgctatcc taatgtatgt ctctatgagt gtatttagcc acacatctgc 2820
ccttggttga ctttctgact cattgcttgc ttgcttgttt ccttgctttg gaaaactatt 2880
gaagattgct aaaaaatacc actgcaaagt gatggaaaag ggtggagaac aggggagtag 2940
ccaggctgga tggctcaaat ataaatgaat gaggaattct ttatgaagta tcagtcagat 3000
tttatgatta agtgatgtaa tataggaatt atgtaaaagg gaagaatgtc tgatactgat 3060
ctattagaga ggtacttttag aggcttcttg attggcataa agttcctaag gttatagatt 3120
ttccccctt ttggctgtat agcaaagtgt tttaatccac gggtgtgcct tattgttcca 3180
ttaaaattgt atcttcgatc catcaataaa tacttgtggt tgaaacaaaa aaaaaaaaaa 3240
aaaaaaaaa 3248

```

<210> 53

<211> 3098

<212> DNA

<213> NM\_003671.2| Homo sapiens CDC14 cell division cycle 14 homolog B (S. cerevisiae) (CDC14B), transcript variant 1, mRNA

```

<400> 53
cacggaacag ccctcctggg gtccccacga gccgcgtcct gctgtgcccc ggcgcctacg 60
cagcagcggc cgcggccgcg gtgggcacgc acggttacc cgggcagctc cggccgccag 120
ctgcagcccc gtcgcctcgg ccgcgccagc cggctgcggg cacctggggg cgggctgggg 180
gcgccggccg cggcaggagg cgctgtagcg agggctgcgg cgccggtcct gcggcggccg 240
cgggaggcag cggggcaggc gctgtggggc gggctcctcc tccggctcct gcgcgaccgc 300
ctcccgccgg gctctgccgg cgcccgccgt ccccgacgcg ccgctctgcg cccgccgccc 360
cgagcgcccc cgcggggctg gcgggagcct cggcgggcgc gcgggcgcgc ggggccatgg 420
tcgtggcccc ctgacgggcc gcggccgcct ccatgaagcg gaaaagcgag cggcggtcga 480

```

gctgggcccgc	cgcgccccc	tgtctcgggc	gctgctcgtc	gacctcgccg	ggtgtgaaga	540
agatccgcag	ctccacgcag	caagacccgc	gccgcggga	ccccaggac	gacgtgtacc	600
tggacatcac	cgatcgctt	tgttttgcca	ttctctacag	cagaccaaag	agtgcacaa	660
atgtacatta	tttcagcata	gataatgaac	ttgaatatga	gaactttctac	gcagattttg	720
gaccactcaa	tctggcaatg	gtttacagat	attgttgcaa	gatcaataag	aaattaaagt	780
ccattacaat	gttaaggaag	aaaattgttc	attttactgg	ctctgatcag	agaaaacaag	840
caaatgctgc	cttccttggt	ggatgctaca	tggttatata	tttggggaga	acccagaag	900
aagcatatag	aatattaatc	tttgagaga	catcctatat	tcctttcaga	gatgctgcct	960
atggaagttg	caattttctac	attacacttc	ttgactgttt	tcatgcagta	aagaaggcaa	1020
tgcagtatgg	cttccttaat	ttcaactcat	ttaaccttga	tgaatatgaa	cactatgaaa	1080
aagcagaaaa	tggagattta	aattggataa	taccagaccg	atttattgcc	ttctgtggac	1140
ctcattcaag	agccagactt	gaaagtgggt	accaccaaca	ttctcctgag	acttatattc	1200
aatattttaa	gaatcacaat	gttactacca	ttattcgtct	gaataaaagg	atgtatgatg	1260
ccaaacgctt	tacggatgct	ggcttcgatc	accatgatct	tttctttgct	gatggcagca	1320
cccctactga	tgccattgtc	aaagaattcc	tagatatctg	tgaaaatgct	gagggtgccca	1380
ttgcagtaca	ttgcaaagct	ggccttggtc	gcacgggcac	tctgatagcc	tgctacatca	1440
tgaagcatta	caggatgaca	gcagccgaga	ccattgcgtg	ggtcaggatc	tgcacacctg	1500
gctcggtgat	tgggcctcag	cagcagtttt	tggatgatgaa	gcaaaccaac	ctctggctgg	1560
aaggggacta	ttttcgtcag	aagttaaagg	ggcaggagaa	tggacaacac	agagcagcct	1620
tctccaaact	tctctctggc	gttgatgaca	ttccataaa	tggggctcag	aatcaagatc	1680
agcaagaacc	cgaaccgtac	agtgatgatg	acgaaatcaa	tggagtgaca	caagggtgata	1740
gacttcgggc	cttgaaaagc	agaagacaat	ccaaaacaaa	cgctattcct	ctcactctct	1800
ccattttcaag	gactaaaaca	gtcttgcggt	aagtaaaaac	ctgtgaccag	agctgaagga	1860
agactctagg	actgaaaact	gcaacagaaa	ttagcacaaat	ttgaaaacaa	aacaaaattg	1920
caaaagcctt	agttgctttt	tccacctaa	aagttgatca	atggagaaaa	tgtccactgg	1980
agtttgaata	atgaactttg	agtttggtg	caagcaaatg	actcagagaa	gggtccagct	2040
ctcaagctga	atgacaaaaca	tgtgtttgta	aatttagtct	cagggtgtaa	tacccaagcc	2100
ctctggtacc	caggagctg	gctggtctgt	ggtgcatgtg	tgtccctgtg	atggcaatca	2160
ttgtagtgtc	tggccttcag	aagaattgag	gatctgatgg	aggtttttta	tgtattttatt	2220
ttctgttcac	cttgtgacct	tgtgtcaaaa	tttataaaga	tacaaaaggc	attactgaaa	2280
tggacttttc	tgtaatgtga	tactatttgg	cttaatcatc	ttcacttgac	tattttgtaat	2340
actgttgtaa	tgttaactct	gttaagtacc	caagctgctt	gtcttccacc	aaagagtgtc	2400
ttattaacaa	gaatctgtga	aaatcacatt	taaacactgt	tgcattgtgt	aagaccaggt	2460
ggtaccttag	taacctaaaa	cttgcaagag	aatattaatg	gtagcttttag	aagactcagg	2520

aggagaaact gacttcagag ttggaagatg ttgcaagtcg ttcccttttc tgccttcag 2580  
 ggactgaaga actgggagggc tgcccattgt ttggttgcca gtcatacaaa ttaaaatcat 2640  
 atttccttcc atgaatggaa gaaacacact attggttttt ccccttgga acagcaatcc 2700  
 caaataatgt cggcttacaa aaaaaaaaaag ttaccacttt tttagagtcc ttccctgtaa 2760  
 cattggattt tttttttccc ttatgagatc cacctaaggc cattgacgtg gcctgcgatc 2820  
 tcagtgacaa tgatctgctt ctggatctca ctgttgccct tggttaggga acacagagtg 2880  
 cttctcccg cagccctactg gaacacagca gagtctgtgc catgaagcag ttacagaaac 2940  
 agaattgatg tgctgctaaa aaaaaaaaaa aaaatggggc ccgggggggc gtccgccggc 3000  
 cctgcgggcc gccgggtgaaa taccactact ctgatcggtt tttcactgac ccggtgaggc 3060  
 gggggggcga gccccgaggg gctctcgctt ctggcgcg 3098

<210> 54

<211> 7850

<212> DNA

<213> XM\_372063.2| PREDICTED: Homo sapiens similar to epiplakin (LOC389697), mRNA

<400> 54  
 atggcagcca cgctgggagc cggcacgccc cccaggcccc aggccaggag catagctggg 60  
 gtgtatgtgg aggcctcggg ccaggcccag agtgtctacg ccgccatgga gcagggcctc 120  
 ctgcctgctg ggctcgggca ggctctgcta gaggcccagg cagccactgg gggcctgggtg 180  
 gacctgccc ggggccagct gctccctgtg tccaaggccc tgcagcaggg tctggtgggg 240  
 ctggagctga aggagaagct gctggccgct gagcgtgcca ctacgggcta tcctgacccc 300  
 tacggcggtg agaagctggc cctctttcag gccatcggga aggagggtgt ggacagggcc 360  
 ctggggcaga gctggctgga ggtccaactg gccactgggg gcctgggtgga cccgcccag 420  
 ggagtgctcg tggccctga gccagcctgc caccagggcc tcctggaccg ggagacatgg 480  
 cacaagctgt cagagcttga gcctggcaca ggtgacctgc gcttcctcga cccaacacg 540  
 ctggagcggc tgacatacca ccagctgctg gaaagggtgtg tgcgtgcccc cggctcgggg 600  
 ctagccttgc tgcccccaaa gatcaccttc cgctccatgg gcggggcggt gaggtcagct 660  
 gagctgctgg aggtgggcat cctggacgag caggctgtgc agggctctcg ggagggcagg 720  
 ctggccgcag tggacgtgag tgcacgtgcc gaggtgcggc gctacctgga gggtagcggc 780  
 agcgtggccg gggttgtcct gctgcccga ggcacaaga agagcttttt ccaggctgcc 840  
 accgagcacc tgctcccaat gggcaccgag ctgccactcc tagaggccca ggctgccacc 900  
 cacaccctgg tggaccccat cacaggccag cggctgtggg tagacgaggc agtcagggcg 960  
 ggccctggtca gccagagct ccatgagcag ctctggtgg cagagcaggc cgtgacaggg 1020



caccacgacc	ccttcagtgg	ctcccaaate	ccccttttcc	aggccatgaa	gaaggggcta	1080
gtggacaggc	caactggcact	gcggctcttg	gatgcccagc	tggccacagg	cgggctggtc	1140
tgtccagcac	gcaggctccg	gctgcccctg	gaggccgccc	tgcgctgcgg	ctgcctggat	1200
gaagacactc	agcggcagct	ctcgagggt	ggcagcttct	cagacggcac	gcacggcggc	1260
ctgcgctatg	aacagctgct	ggccctctgt	gtcaccgacc	cagagaccgg	gcttgccttc	1320
ctgccactct	caggggggacc	ccggggagg	gagccccagg	gacccccatt	catcaagtac	1380
agcactcggc	aggccctgag	cacggccaca	gccaccgtct	ctgtggggaa	gttccggggc	1440
cggcccgtgt	ccctctggga	gctgctcttc	tctgaggcca	tctcctcaga	gcagagggcg	1500
atgctggccc	agcagtacca	ggaagggacc	ctctccgtgg	agaagctggc	cgctaagctg	1560
agcgccaccc	tcgagcaggc	tgcagccact	gccagggtca	ccctttctgg	gctgaggggac	1620
accgtgacac	caggagagct	gctgaaagcc	gagatcatcg	accaggacct	gtacgagcgg	1680
ctggagcatg	gacaggccac	agccaaggat	gtgggcagcc	tggcctcggg	gcagaggtac	1740
ctgcagggtg	cgggctgcat	tgctggcctg	ctgctccctg	gctcccagga	acgcctgagc	1800
atctatgagg	cccgatgcaa	ggggctcctc	cggcccggca	ctgccctcat	ccttctggag	1860
gcacaagctg	ccacaggctt	catcatcgac	ccaaaagcaa	acaaggggca	ctccgttgag	1920
gaggcactga	gggctgctgt	cattgggcct	gatgtgttcg	cgaagctgct	gtcggctgag	1980
cgcgctgtca	ctggctacac	tgacccttac	accgggcagc	agatctccct	cttccaggcc	2040
atgcagaagg	gcctcatcgt	ccgggagcac	ggcatccgcc	tgctggaggc	ccagatcgcc	2100
acgggcggcg	tcatcgaccc	cgtgcacagc	caccgcgtgc	ccgtggacgt	ggcctaccgg	2160
cgcgctact	tcgatcagat	gctgaacttg	atcctgttgg	acccttctga	cgacaccaag	2220
ggcttcttcg	accccaacac	gcacgagaac	ctcacgtacc	tgcagcttct	ggagcgctgt	2280
gtgcgtgacc	ccgagacggg	cctgtacctc	ctgccactca	gcagcacgca	gtccccgctg	2340
gtggacagtg	ccaccagca	ggccttcag	aacctgctgc	tctccgtgaa	gtatggacgg	2400
tttcaggggc	agaggggtctc	cgcggtggag	ctgatcaact	ctgagtactt	cagcgagggc	2460
cgcaggaggc	agctgctgcg	tcgctaccgg	cagcgcgagg	tcacgctggg	gcaggtggca	2520
aagctgctgg	aggcggagac	gcagagacag	gcggacatca	tgctgcccgc	actgcggagc	2580
cgggtcaccg	tccaccagct	cctggaggcc	ggtatcattg	accagcagct	gttggaccaa	2640
gtgctggccg	ggacaatcag	cccggaggcc	ctcctactca	tggacggcgt	ccgcaggtag	2700
ctgtgcggcc	tgggagctgt	gggcgggtgtg	cggtgctgc	cctctggcca	gcggctcagc	2760
ctctaccagg	ccatgaggca	gaagctgctg	gggcccagg	tggccctggc	cctgctggag	2820
gcccaggcgg	ccaccggaac	catcatggac	cctcacagcc	cagagagcct	ctcgggtgat	2880
gaggccgtgc	gcaggggtgt	ggtggggccg	gagctgtatg	gcaggctgaa	gcgggctgag	2940
ggtgccattg	ctggcttcag	agacccttc	tctgggaagc	aggtgtctgt	gttccaggcc	3000

atgaagaaag	gtctcatccc	ttgggagcaa	gctgcccgcc	tcctggaggc	tcaagtggcc	3060
acaggagggga	tcattgaccc	caccagccac	caccacctcc	ccatgccagt	ggccattcag	3120
cgtggctatg	ttgaccagga	gatggagaca	gccttggtcca	gtcctccga	aaccttcccc	3180
acaccggacg	gccaggggcg	cacgagctat	gcccagctcc	tggaggagtg	ccccagggat	3240
gagacttctg	gccttcacct	cctgcccctg	ccagaaagtg	ctcctgccct	ccccaccgag	3300
gagcaggtcc	agaggagcct	gcaggccgtg	ccggggggcca	aggatggcac	atccctctgg	3360
gacctgctca	gtccttgcca	cttcaccgag	gagcaacgga	ggggcctgct	ggaggacgtg	3420
caggagggga	ggaccactgt	gccacagctg	ctagcctctg	tgcagagggtg	ggtacaggag	3480
accaagctcc	tggcccaggc	ccgcgtcatg	gtgcccggcc	cacggggtga	ggtacccgct	3540
gtctggctgc	tggatgctgg	catcatcacc	caggagaccc	ttgaggccct	ggctcagggc	3600
acgcagtcgc	ccgcccagggt	cgccgagcag	ccggcggtga	aggcctgcct	gtggggcaca	3660
ggctgcgtgg	ccggtgtgct	gctacagccc	tctggggcca	aggccagcat	cgcccaggcc	3720
gtgagggatg	gcctcctgcc	cacaggcctg	ggccagaggc	tgctggaagc	ccagggtggca	3780
tctggcttcc	ttgttgaccc	cctgaacaac	cagagactgt	cagtggagga	cgcggttaag	3840
gtcggcctgg	tgggcaggga	gctgagtgag	cagctcgggc	aggccgagag	ggcggcggcc	3900
gggtaccag	atccctactc	tagggcctcc	ctctctctgt	ggcaggccat	ggagaagggg	3960
ctcgtgccac	agaacgaggg	cttgcccctc	ctgcaggtgc	agctggccac	aggggggtgtg	4020
gtggaccctg	tccacgggggt	gcacctgccc	caggcggcag	cctgcagact	cgcccttctg	4080
gacacacaga	cgagccagggt	gctgactgca	gttgacaagg	acaacaagtt	cttctttgac	4140
cccagtgcg	gggaccagggt	gacctaccag	cagctcaggg	agcgctgctg	gtgctgactcc	4200
gagaccggat	tggtgtgtgt	gccactgccc	tcagacacag	tgcttgaggt	ggacgaccac	4260
accgcgggtg	ctctgagggc	catgaagggtg	cccgtcagca	cagggaggtt	taaggggtgt	4320
agcgtgtcac	tctgggacct	gctgctctcc	gaatacgttg	gcgctgacaa	gcggcggggag	4380
ctggtggcac	tctgtcggtc	tgggagggct	gcggccctgc	ggcaggtggt	cagcgcagtc	4440
accaccctgg	tcgaggctgc	agagaggcag	cccctgcagg	ccaccttcag	agggctccgg	4500
aagcaggtgt	cagccaggga	cctgttcagg	gcgcagctga	tcagcaggaa	gacgctggac	4560
gagctgagcc	aggggacaa	gactgtgaag	gaggtggcgg	agatggacag	cgtgaagcgg	4620
tccctggagg	gaggcaactt	cattgccggg	gtccttatcc	agggcaccca	ggagaggatg	4680
agcatcccag	aggccctgag	gaggcacatc	ctgcggcctg	gcacagccct	ggtgctgctg	4740
gaggcacagg	cagctaccgg	cttcatcatc	gaccccgctg	agaaccggaa	gctgaccgtg	4800
gaggaggcgt	tcaaagcagg	aatgttcggg	aaagaaacct	acgtgaagct	gctgtcggcc	4860
gagcgcgccg	tcaccggcta	caccgacccc	tataccgggc	agcagatctc	cctcttccag	4920
gccatgcaga	aggacctcat	cgtccgggag	cacggcatcc	gcctgctgga	ggcccagatc	4980
gccacgggcg	gcatcatcga	ccccgtgcac	agccaccgcg	tgcccgtgga	cgtggcctac	5040

cgctgctggct	acttcgacga	ggagatgaac	cgcatcctgg	cggaccccag	cgacgacacc	5100
aagggcttct	tcgaccccaa	cacgcacgag	aacctcacgt	acctgcagct	tctggagcgc	5160
tgtgtggagg	accccagagc	gggcctgtac	ctgctacaaa	tcataaagaa	aggagaaaac	5220
tacgtgtaca	tcaatgaggg	cacgagacac	gtgttgcaat	ccagaactgc	aaaaatgcgc	5280
gtggggaggt	ttgctgacca	ggtggtctct	ttctgggacc	tgctgtcctc	tccatacttc	5340
acagaggaca	ggaagcggga	gctcatccag	gagtatggag	cccagagtgg	gggcctggag	5400
aaattgctgg	aaatcatcac	cacgacaatt	gaagaaacag	agacgcaaaa	ccaaggcatc	5460
aaagtggcgg	ccatcagagg	ggaggtgaca	gctgcagacc	tgttcaactc	caggggtcatc	5520
gatcagaaga	ccctgcacac	acttcgtgtg	gggaggactg	ggggacaggc	actcagcacg	5580
ctggagtgtg	tgaagcccta	tctggaaggc	agcggctgca	ttgcgggggt	cacggtgccc	5640
tccaccaggg	aggtcatgag	cctccatgag	gccagcagga	aggagctcat	ccctgcagca	5700
tttgcgactt	ggctgctgga	ggcgaggcc	gccaccgggt	tcctcctgga	cccctgcacc	5760
cgccagaagc	tctctgtgga	tgaggctgtg	gatgtgggccc	tggtgaacga	ggagctgcgg	5820
gagaggctcc	tgaaggctga	aagagctgcc	acgggctaca	gggatccggc	cacaggagac	5880
acgatcccgc	tgttccaggc	catgcagaag	cagctcatcg	agaaggcgga	ggcactgagg	5940
ctgctggagg	tgagggtggc	cacgggggggt	gtcatcgacc	cacagcacca	ccaccggctc	6000
ccactggaaa	cagcctacag	acggggctgt	ctgcacaagg	acatctatgc	gctcatttcc	6060
gaccagaagc	acatgaggaa	acggtttgtg	gacccgaaca	cgcaagagaa	ggtctcgtac	6120
cgagagctgc	aggagagggtg	ccgcccacaa	gaggacacgg	gctggctgct	gttcccagtg	6180
aacaaggctg	cacgggactc	cgagcacatc	gatgacgaga	cgagaagggc	cctggaggca	6240
gagcaagtgg	aaatcacagt	gggaagggtt	agaggccaga	aaccaacact	gtgggcacta	6300
ctgaattccg	aatacgtgac	agaggagaag	aagctccagc	tggtgaggat	gtatagaaca	6360
cacaccagac	gggcactgca	gacggtagcg	cagctcatct	tagagttgat	cgagaagcag	6420
gaaaccagca	acaaacacct	gtggttccaa	ggaattagac	gacagatcac	agcttctgaa	6480
ctcctcagct	cagccataat	cacggaggaa	atgctccagg	acctggaaac	gggacggagc	6540
acgacgcaag	agctcatgga	ggacgaccgc	gtcaagcgct	acctggaggg	caccagctgc	6600
atcgcgggcg	tcctggtgcc	cgccaaggac	cagcccggcc	gccaggagaa	gatgagcatc	6660
taccaggcca	tgtggaaggg	cgtgctgcgg	cccggcacgg	ccctggtgct	gctggaggcg	6720
caggcgccca	ccggcttcgt	catcgacccc	gtgcgcaacc	tgaggctgtc	ggtggaggag	6780
gccgtggctg	cgggcgtggg	gggcggcgag	atccaggaga	agctgctgtc	ggccgagcgc	6840
gccgtcaccg	gctacaccga	cccctacacc	gggcagcaga	tctccctctt	ccaggccatg	6900
cagaaggacc	tcatcgtccg	ggagcacggc	atccgcctgc	tgaggggcca	gatcgccacg	6960
ggcggcgctc	tcgaccccg	gcacagccac	cgcgtgcccc	tggaactggc	ctaccggcgc	7020

```

ggctacttcg acgaggagat gaaccgcgtc ctggccgacc ccagcgacga caccaagggc 7080
ttcttcgacc ccaacacgca cgagaacctc acgtacctgc agcttctgca gagggccacc 7140
ctggaccctg agacggggct cctatttctt tctctctctc tacagtgact gggcttcctc 7200
cgtgcagttt tctgcaactc tggagaagtg gaggcatact tgttgtctctg ggttggtttt 7260
tttttgtttt ttttttgtc attctttaat tttgttgttt taccattcg ttatctgtgg 7320
aaaacgtttt aagtgtgcat gtgacagaaa ctttcccttt gtccatcgag gtgtttcata 7380
agttttttgg tgtgttttct gggtcgtcta tgtgtcatat ggttttactt ttctctcctt 7440
tttcgttttc agaactttt tctgtctgtt ttggattcac tgcttcatt ttacagaatg 7500
tcactcttta gactctcagt ccatcatgcc atcgggtact cttgttgag tgtaattttt 7560
attacatgcg gttatttccc taacgatgtg ctattcacgt tcattctcaa actcattttc 7620
catcagccag tgtctactat ttagtgccct ggctctatct cggctcctcct ccccgggctt 7680
tccctggctg ctgtgctggc caaaagcatg ggctttattc tctccattgg ctgctgctcc 7740
accttagagg tgtgacctca ctagcgttga ctgagcgagt ctgttggtga gaagaacttt 7800
ttgtagtaat ttactaggaa aaattctgaa caagtaaaat atgaaggaaa 7850

```

<210> 55

<211> 454

<212> DNA

<213> NM\_005952.2| Homo sapiens metallothionein 1X (MT1X), mRNA

```

<400> 55
tctgtcccg cgcgtgtttt cctcttgatc gggaactcct gcttctcctt gcctcgaaat 60
ggaccccaac tgctcctgct cgcctgttgg ctctgtgccc tgtgccggct cctgcaaagt 120
caaagagtgc aaatgcacct cctgcaagaa gagctgctgc tcctgctgcc ctgtgggctg 180
tgccaagtgt gccaggggct gcatctgcaa agggacgtca gacaagtgca gctgctgtgc 240
ctgatgccag gacagctgtg ctctcagatg taaatagagc aacctatata aacctggatt 300
ttttttttt ttttttgta caaccctgac ccgtttgcta catctttttt tctatgaaat 360
atgtgaatgg caataaattc atctagacta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 420
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 454

```

<210> 56

<211> 2090

<212> DNA

<213> NM\_003242.3| Homo sapiens transforming growth factor, beta receptor II (70/80kDa) (TGFB2), mRNA

<400> 56

gttggcgagg agtttctgt ttccccgcga gcgctgagtt gaagttgagt gagtcaactcg	60
cgcgcacgga gcgacgacac ccccgcgcggt gcacccgctc gggacaggag ccggactcct	120
gtgcagcttc cctcggccgc cgggggcctc cccgcgctc gccggcctcc agggccctcc	180
tggctggcga gcgggcgcca catctggccc gcacatctgc gctgccggcc cggcgcgggg	240
tccggagagg gcgcggcgcg gagcgcagcc aggggtccgg gaaggcgccg tccgtgcgct	300
gggggctcgg tctatgacga gcagcggggt ctgccatggg tcgggggctg ctcaggggcc	360
tgtggccgct gcacatcgtc ctgtggacgc gtatcgccag cacgatccca ccgcacgttc	420
agaagtcggt taataacgac atgatatgca ctgacaacaa cgggtgcagtc aagtttccac	480
aactgtgtaa attttgtgat gtgagatttt ccacctgtga caaccagaaa tcttgcata	540
gcaactgcag catcacctcc atctgtgaga agccacagga agtctgtgtg gctgtatgga	600
gaaagaatga cgagaacata aactagaga cagtttgcca tgaccccaag ctcccctacc	660
atgactttat tctggaagat gctgcttctc caaagtgcac tatgaaggaa aaaaaaagc	720
ctggtgagac tttcttcatg tgttctctga gctctgatga gtgcaatgac aacatcatct	780
tctcagaaga atataacacc agcaatcctg acttggtgct agtcataatt caagtgcag	840
gcacagcct cctgccacca ctgggagttg ccatactgt catcatcatc ttctactgct	900
accgcgttaa ccggcagcag aagctgagtt caacctggga aaccggcaag acgcggaagc	960
tcatggagtt cagcgagcac tgtgccatca tctggaaga tgaccgctct gacatcagct	1020
ccacgtgtgc caacaacatc aaccacaaca cagagctgct gccattgag ctggacaccc	1080
tgggtggggaa aggtcgcttt gctgaggtct ataaggccaa gctgaagcag aacacttcag	1140
agcagtttga gacagtggca gtcaagatct ttccctatga ggagtatgcc tcttgggaaga	1200
cagagaagga catcttctca gacatcaatc tgaagcatga gaacatactc cagttcctga	1260
cggctgagga gcggaagacg gagttgggga aacaatactg gctgatcacc gccttcacg	1320
ccaagggcaa cctacaggag tacctgacgc ggcagtgcac cagctgggag gacctgcgca	1380
agctgggcag ctccctcgcc cgggggattg ctcacctcca cagtgatcac actccatgtg	1440
ggaggcccaa gatgcccac gtgcacaggg acctcaagag ctccaatatc ctcgtgaaga	1500
acgacctaac ctgctgcctg tgtgactttg ggctttccct gcgtctggac cctactctgt	1560
ctgtggatga cctggctaac agtgggcagg tgggaactgc aagatacatg gctccagaag	1620
tcctagaatc caggatgaat ttggagaatg ctgagtcctt caagcagacc gatgtctact	1680
ccatggctct ggtgctctgg gaaatgacat ctcgctgtaa tgacgtggga gaagtaaaag	1740
attatgagcc tccatttgggt tccaagggtgc gggagcacc ctgtgtcgaa agcatgaagg	1800
acaacgtgtt gagagatcga gggcgaccag aaattcccag cttctggctc aaccaccagg	1860
gcatccagat ggtgtgtgag acgttgactg agtgctggga ccacgaccca gagggccgctc	1920

tcacagccca	gtgtgtggca	gaacgcttca	gtgagctgga	gcatctggac	aggctctcgg	1980
ggaggagctg	ctcggaggag	aagattcctg	aagacggctc	cctaaacact	accaaatagc	2040
tcttatgggg	caggctgggc	atgtccaaag	aggctgcccc	tctcaccaaa		2090

<210> 57

<211> 4568

<212> DNA

<213> NM\_012408.3| Homo sapiens protein kinase C binding protein 1 (PRKCBP1), transcript variant 2, mRNA

<400> 57						
gtgagaacta	ggagcctgtc	ctccatgttt	tataagtatt	gacattacac	agtgttaaca	60
atgcatccac	agagcttggc	tgaagaggaa	ataaaaacag	aacaggaggt	ggtagagggc	120
atggatatct	ctactcgtc	caaagatcct	ggctctgcag	agagaacagc	ccagaaaaga	180
aagttcccca	gccctccaca	ttcttccaat	ggccactcgc	cgcaggacac	atcaacaagc	240
cccattaaaa	agaaaaagaa	acctggctta	ctgaacagta	acaataagga	gcagtcagaa	300
ctaagacatg	gtccgtttta	ctatatgaag	cagccactca	ccacagaccc	tgttgatggt	360
gtaccgcagg	atggacggaa	tgattttctac	tgctgggttt	gtcaccggga	aggccaagtc	420
ctttgctgtg	agctctgtcc	ccgggtttat	cacgctaagt	gtctgagact	gacatcggaa	480
ccagaggggg	actggttttg	tcctgaatgt	gagaaaatta	cagtagcaga	atgcatcgag	540
acccagagta	aagccatgac	aatgtctacc	attgaacagt	tatcctacct	gctcaagttt	600
gccattcaga	aaatgaaaca	gccagggaca	gatgcattcc	agaagcccgt	tccattggaa	660
cagcaccctg	actatgcgga	atacatcttc	catccaatgg	acctttgtac	attggaaaag	720
aatgcgaaaa	agaaaatgta	tggtcgcaca	gaagccttcc	tggtgatgc	aaagtggatt	780
ttgcacaact	gcatcattta	taatggggga	aatcacaaat	tgacgcaaat	agcgaaaagta	840
gtcatcaaaa	tctgtgaaca	tgagatgaat	gaaatcgaag	tatgtccaga	atgttatcta	900
gctgcttgcc	aaaaacgaga	taactggttt	tgtgagcctt	gtagcaatcc	acatcctttg	960
gtctggggcca	aactgaaggg	gtttccattc	tggcctgcaa	aagctctaag	ggataaagac	1020
gggcaggctg	atgcccgatt	ctttggacaa	catgacaggg	cctgggttcc	aataaataat	1080
tgctacctca	tgtctaaaga	aattcctttt	tctgtgaaaa	agactaagag	catcttcaac	1140
agtgccatgc	aagagatgga	ggtttacgtg	gagaacatcc	gcaggaagtt	tgggggtttt	1200
aattactctc	catttaggac	accctacaca	cccaacagcc	agtatcaaat	gctgctcgat	1260
cccaccaacc	ccagcgccgg	cactgccaag	atagacaagc	aggagaaggt	caagctcaac	1320
tttgacatga	cggcatcccc	caagatcctg	atgagcaagc	ctgtgctgag	tgggggcaca	1380
ggccgcccga	tttccttgtc	ggatatgccg	cgctcccca	tgagcacaaa	ctcttctgtg	1440

cacacgggct	ccgacgtgga	gcaggatgct	gagaagaagg	ccacgtcgag	ccacttcagt	1500
gcgagcgagg	agtccatgga	cttcctggat	aagagcacag	cttcaccagc	ctccaccaag	1560
acgggacaag	cagggagttt	atccggcagc	ccaaagccct	tctctcctca	actgtcagct	1620
cctatcacga	cgaaaacgga	caaaacctcc	accaccggca	gcacacctga	tcttaacctg	1680
gatcgaagca	aagctgagat	ggatttgaag	gagctgagcg	agtcgggtcca	gcaacagtcc	1740
acccctgttc	ctctcatctc	tcccaagcgc	cagattcgta	gcaggttcca	gctgaatctt	1800
gacaagacca	tagagagttg	caaagcacia	ttaggcataa	atgaaatctc	ggaagatgtc	1860
tatacggccg	tagagcacag	cgattcggag	gattctgaga	agtcagatag	tagcgatagt	1920
gagtatatca	gtgatgatga	gcagaagtct	aagaacgagc	cagaagacac	agaggacaaa	1980
gaagggtgtc	agatggacaa	agagccatct	gctgttaaaa	aaaagcccaa	gcctacaaac	2040
ccagtggaga	ttaaagagga	gctgaaaagc	acgtcaccag	ccagcgagaa	ggcagaccct	2100
ggagcagtca	aggacaaggc	cagccctgag	cctgagaagg	acttttccga	aaaggcaaaa	2160
ccttcacctc	accccataaa	ggataaactg	aagggaaaag	atgagacgga	ttccccaaca	2220
gtccatttgg	gcctggactc	tgattcagag	agcgaacttg	tcatagattt	aggagaagac	2280
cattctgggc	gggagggtcg	aaaaataaag	aaggaaccca	aagaaccatc	tcccaaacag	2340
gatgttgtag	gtaaaactcc	accatccacg	acggtgggca	gccattctcc	cccggaaaca	2400
ccggtgctca	cccgtcttcc	cgcccaaact	tccgcggtcg	gcgccacagc	caccaccagc	2460
acgtcctcca	cggtcacctg	cacggccccg	gcccccgccg	ccacaggaag	cccagtgaaa	2520
aagcagaggg	cgctttttacc	gaaggagact	gccccggccg	tgcagcgggt	cgtgtggaac	2580
tcatcaactg	tccagcagaa	ggagatcaca	cagagcccat	ccacgtccac	catcaccttg	2640
gtgaccagca	cacagtcac	gccccctggtc	accagctcgg	ggtccatgag	cacccttggtg	2700
tcctcagtca	acgttgacct	gcccacgcgc	actgcctcag	ctgatgtcgc	cgctgatatt	2760
gccaagtaca	ctagcaaaat	gatggatgca	ataaaaggaa	caatgacaga	aatatacaac	2820
gatctttcta	aaaacactac	tggaagcaca	atagctgaga	ttcgcaggct	gaggatcgag	2880
atagagaagc	tccagtggct	gcaccagcaa	gagctctccg	aaatgaaaca	caacttagag	2940
ctgaccatgg	cggagatgcg	gcagagcctg	gagcaggagc	gggaccggct	catcgccgag	3000
gtgaagaagc	agctggagtt	ggagaagcag	caggcggtgg	atgagaccaa	gaagaagcag	3060
tggtgcgcca	actgcaagaa	ggaggccatc	ttttactgct	gttggaacac	tagctactgt	3120
gactacccct	gccagcaagc	ccactggcct	gagcacatga	agtcctgcac	ccagtacagt	3180
actgctcctc	agcaggaagc	ggatgctgag	gtgaacacag	aaacactaaa	taagtccctc	3240
caggggagct	cctcgagcac	acaatcagca	ccttcagaaa	cggccagcgc	ctccaaagag	3300
aaggagacgt	cagctgagaa	aagcaaggag	agtggctcga	cccttgacct	ttctggctcc	3360
agagagacgc	cctcctccat	tctcttaggc	tccaaccaag	gctctgacca	ttcccggagt	3420
aataaatcca	gttgagcag	cagtgatgag	aagaggggat	cgacacgttc	cgatcacaac	3480

accagtacca gcacgaagag cctcctcccg aaagagtctc ggctggacac cttctgggac 3540  
 tagcagtga tctgggacaca aaccaccac cccattggga gaaaaaccca gacgccagga 3600  
 aaagaagaaa caacaaaggc aggagaacag ccactttcag acttgaaaat gacaaaaccc 3660  
 tcagttgagc ctgagccccc ggcgcggggg ctgctacact acaggacacc cagcatcggc 3720  
 tttgactgca gactgttcac ccacacgagc cctgtgcttt tgggtgtaa atgtacaat 3780  
 ttgtggatgt cattgaatct agaggacttt ccccttttta ttttgtatt aactttaact 3840  
 tattaataaaa aaaaaaagaa aaagaaaaac gatttaaaaa aaaaaaaaaa agcaaccaac 3900  
 cccaacaaca aaaaagaatg ttttgggtatt ggagaaggga tggtcagtta gcctgtctgt 3960  
 cacacgacgg aatggatact gggcccgggg accactttca tactcacgtc ctcaccttg 4020  
 gatacccagg ggagggcgaa ccgttttcgc tctgtgtct gtacgcagca tgttgggac 4080  
 gggagtttcg gcacagacta tcccatcaag ccgttggctc ctttcagcta ctacgttacc 4140  
 acgttcctaa aacgcaagct ctccggacca gacggacaca gggagaagct agtttctttc 4200  
 atgtgattga aatgatgact ctactcctaa aagggaataa acaatatact tgtttacaga 4260  
 agagaaacaa acaagcccca ctacgctcag tcacaggaga gaacacagaa agtcttagga 4320  
 tcatgaactc tgaaaaaaag agaaacctta tctttgcttt gtggttcctt taaacacact 4380  
 cacacacact tggtcagaga tgctgtgctt cttggaagca aggactcaaa ggcaagggtgc 4440  
 acgcagagga cgtttgagtc tgggatgaag catgtacgta ttatttatat gatggaattt 4500  
 cacgttttta tgtaagcatg aaacacaggc agtatgagag aaagcaaggc ccgtcatgct 4560  
 gtccgtac 4568

<210> 58

<211> 2069

<212> DNA

<213> NM\_003270.2| Homo sapiens transmembrane 4 superfamily member 6 (TM4SF6), mRNA

<400> 58  
 cgctcgtaag ttttcggcag tttccgggga gactcgggga ctccgcgtct cgctctctgt 60  
 gttccaatcg cccggtgcgg tgggtgcaggg tctcgggcta gtcatggcgt ccccgctctg 120  
 gagactgcag actaaaccag tcattacttg tttcaagagc gttctgctaa tctacacttt 180  
 tattttctgg atcactggcg ttatccttct tgcagttggc atttggggca aggtgagcct 240  
 ggagaattac ttttctcttt taaatgagaa ggccaccaat gtccccttcg tgctcattgc 300  
 tactggtacc gtcattattc ttttgggcac ctttgggtgt tttgctacct gccgagcttc 360  
 tgcatggatg ctaaaactgt atgcaatgtt tctgactctc gtttttttgg tcgaactggg 420  
 cgctgccatc gtaggatttg ttttcagaca tgagattaag aacagcttta agaataatta 480



tgagaaggct ttgaagcagt ataactctac aggagattat agaagccatg cagtagacaa	540
gatccaaaat acgttgcatt gttgtggtgt caccgattat agagattgga cagatactaa	600
ttattactca gaaaaaggat ttcctaagag ttgctgtaaa cttgaagatt gtactccaca	660
gagagatgca gacaaagtaa acaatgaagg ttgttttata aaggatgatga ccattataga	720
gtcagaaatg ggagtcgttg caggaatttc ctttggagtt gcttgcttcc aactgattgg	780
aatctttctc gcctactgcc tctctcgtgc cataacaaat aaccagtatg agatagtgtg	840
acccaatgta tctgtgggcc tattcctctc tacctttaag gacatttagg gtccccctg	900
tgaattagaa agttgcttgg ctggagaact gacaacacta cttactgata gaccaaaaaa	960
ctacaccagt aggttgattc aatcaagatg tatgtagacc taaaactaca ccaataggct	1020
gattcaatca agatccgtgc tcgcagtggg ctgattcaat caagatgtat gtttgctatg	1080
ttctaagtcc accttctatc ccattcatgt tagatcgttg aaaccctgta tccctctgaa	1140
acactggaag agctagtaaa ttgtaaatga agtaatactg tgttcctctt gactgttatt	1200
tttcttagta gggggccttt ggaaggcact gtgaatttgc tattttgatg tagtgttaca	1260
agatggaaaa ttgattcctc tgactttgct attgatgtag tgtgatagaa aattcacccc	1320
tctgaactgg ctccctccca gtcaagggtta tctggtttga ttgtataatt tgcaccaaga	1380
agttaaaaatg ttttatgact ctctgttctg ctgacaggca gagagtcaca ttgtgtaatt	1440
taatttcagt cagtcaatag atggcatccc tcatcagggt tgccagatgg tgataacagt	1500
gtaaggcctt ggggtctaagg catccacgac tggaagggac tactgatgtt ctgtgataca	1560
tcaggtttca gcacacaact tacatttctt tgcctccaaa ttgaggcatt tattatgatg	1620
ttcatacttt ccttcttggt tgaaagtttc taattattaa atggtgtcgg aattgttgta	1680
ttttccttag gaattcagtg gaacttatct tcattaaatt tagctggtag caggttgata	1740
tgacttgtca atattatggt caactttaag tcttagtttt cgtttgtgcc ttgattaat	1800
aagtataact cttatacaat aaatactgct ttcctctaaa aagatcgtgt ttaaattaac	1860
ttgtagaaaa tctgctggaa tgggttggtt tttccactga gaaagctaag ccctacattt	1920
ctattcagag tactgttttt agatgtgaaa tataagcctg cggccttaac tctgtattaa	1980
aaaaaatgtt tttgtttaaa aaaaactgtt cccatagggtg cagcaaacca ccatggcaca	2040
tgtataccta tgtaacaaac ctgcacatt	2069

<210> 59

<211> 2402

<212> DNA

<213> NM\_021200.1| Homo sapiens plectstrin homology domain containing, family B (evectins) member 1 (PLEKHB1), mRNA

<400> 59  
aagagaggaa ggcttaaaga gccagactgc gcagccagga ctggggtatg ggcgctgtcc 60  
tgcaggccaa agaatgaaga tgtagccccg cccccaacct aggaggagga ccagcccggg 120  
tcctgtcctg cccccgcaac ctgcctccga ttccactccg ggaacctcgg cgatgctgag 180  
ccaagaccac ttctgaatca gggatgactt gtctagtga cgtaggggtca gagccatcag 240  
ttggaaaggc tgggaggagc ctggagaaag aggcgacctt ccttgggatc tgtgcgctcc 300  
ctccttgccct cccctccag cctccactt ggtagacctt tcctgatccc cttatctcta 360  
aggcgctcag ggaaatgccc cgctgcggga gccttctggg aaatgctgcc ctggccaccc 420  
aggaaccatg agccctgcag ccccggtccc gcctgactcc gctctggaaa gtccttttga 480  
agaaatggcc ctggtgaggg gcggctggct gtggagacag agctccatcc tccgccgctg 540  
gaagcggaaac tggtttgccc tgtggctgga cgggaccctg ggatactacc acgatgagac 600  
agcgcaggac gaggaggacc gtgtgctcat ccacttcaat gtccgtgaca taaagatcgg 660  
cccagagtgc catgatgtgc agccccaga gggccggagc cgagatggcc tgctgactgt 720  
gaacctacgg gaaggcggcc gcctgcacct ctgtgcggag accaaggatg atgccctagc 780  
atggaagaca gcactgctgg aggcaaaact caccctggcc ccagctggag ccaccgtccc 840  
tcccaggagc cgccgggttt gctccaaggt cagggtgtgtg acccgctcgt ggagcccctg 900  
taaggttgag aggcggatct ggggtgcgct ctacagccg taccaagact actacgaggt 960  
ggtgcccccc aatgcacacg aggccacgta tgtccgcagc tactacggac cgccctacgc 1020  
aggccctggc gtgacgcacg tgatagtgcg ggaggatccc tgctacagcg ccggcgcccc 1080  
tctggccatg ggcattgctt cgggagccgc cactggggcg gcgctgggct cgctcatgtg 1140  
gtcgccctgc tggttctgag ccctgggact cggagcactg acccctgcgc ttggattgct 1200  
agactcctct tcctcctgga ccccatctc taccatccaa gccctgtccc actttggccc 1260  
tatcctctcc attagctcct tccgggtttg gaccattccc cccactccct acccttaatc 1320  
cccacatggg aagaagctat catcacaggt acaaacatcg cttgaagtct tcacatctac 1380  
cactagacac ccccaaaatc tgttatagac atttatggat acatttcctc taaacacaac 1440  
agggcacagc aaatacgact tcatttggtc tcgagttccc caggcgctgt agacacaaca 1500  
tgaatcgggc tctctgctct ctcttaggg agctcgagtc ctgggtggga gaacaggagt 1560  
aaacaaggac ttgacaaagc tgaagagtta tcagtccttt gacaaggaca ggtggggcag 1620  
ggagcaagac aggtaggctg gaagaacagt tattggcaag tatgcagagc cgtgaacgtc 1680  
atggcatgtc caaggaatta aatgggagtt catttgggct ggggtggagg ctgggatcag 1740  
accgtgggtg gccttcaagc taaggagctt cctaggtgaa aggggagatg tgagccttct 1800  
ctggagggaa gtttcatgat tgcatttata atgaatatat tgcctgtttt gtgaatactg 1860  
acacatgtcc atacctaaaa cactcctgag ttaagtccca tccttcccac aaacagcttc 1920  
ctggctggta cccatgataa caattgagct gaacctgggg acccctgggt ggggaacagg 1980

tgagttctat ttgagacttc cagccctaga aagctgcctc cgtccagaaa tgcctctcac	2040
accaggagct cggccctctc tttatagctg tgactgtcac cctctcaggc tttgtctcat	2100
ccttcattct gaataagatg gcagtgttct cctctggggc ctgatccacc tctacaccag	2160
cccaggaagc cccatctgtg cctgccctca ggtgggtccac cagtctcccc ctttggttcc	2220
cttcaggtct cttccccctt tctatcccaa tcaccaatag aaatgctaac atccctgcct	2280
ggtagccaga ctagcccact aaagctcccc tgtaaatggg ggctccatta gttctgctgc	2340
cgagactaat aaagatttgg ttggttctag cagtaaaaaa aaaaaaaaaa aaaaaaaaaa	2400
aa	2402

<210> 60

<211> 2856

<212> DNA

<213> NM\_003661.2| Homo sapiens apolipoprotein L, 1 (APOL1), transcript variant 1, mRNA

<400> 60	
actttccctt tcgaattcct cggtatatct tggggactgg aggacctgtc tggttattat	60
acagacgcat aactggagggt gggatccaca cagctcagaa cagctggatc ttgctcagtc	120
tctgccaggg gaagattcct tggaggaggc cctgcagcga catggaggga gctgctttgc	180
tgagagtctc tgcctctgtc atctggatga gtgcactttt ccttgggtgtg ggagtgaggg	240
cagaggaagc tggagcgagg gtgcaacaaa acgttccaag tgggacagat actggagatc	300
ctcaaagtaa gcccctcggg gactgggctg ctggcaccat ggaccagag agcagtatct	360
ttattgagga tgccattaag tatttcaagg aaaaagttag cacacagaat ctgctactcc	420
tgctgactga taatgaggcc tggaacggat tcgtgggtgc tgctgaactg cccaggaatg	480
aggcagatga gctccgtaaa gctctggaca accttgcaag acaaatgatc atgaaagaca	540
aaaactggca cgataaaggc cagcagtaca gaaactgggt tctgaaagag tttcctcggg	600
tgaaaagtga gcttgaggat aacataagaa ggctccgtgc ccttgcagat ggggttcaga	660
aggtccacaa aggaccacc atcgccaatg tgggtgtctg ctctctcagc atttcctctg	720
gcatcctgac cctcgtcggc atgggtctg cacccttcac agaggaggc agccttgtag	780
tcttggaacc tgggatggag ttgggaatca cagccgcttt gaccgggatt accagcagta	840
ccatggacta cggaaagaag tgggtggacac aagcccaagc ccacgacctg gtcataaaaa	900
gccttgacaa attgaaggag gtgagggagt ttttgggtga gaacatatcc aactttcttt	960
ccttagctgg caatacttac caactcacac gaggcattgg gaaggacatc cgtgccctca	1020
gacgagccag agccaatctt cagtcagtac cgcgtgcctc agcctcacgc cccggggtca	1080
ctgagccaat ctcagctgaa agcgggtgaac aggtggagag ggttaatgaa cccagcatcc	1140

tggaaatgag	cagaggagtc	aagctcacgg	atgtggcccc	tgtaagcttc	tttcttgtgc	1200
tggtatgtagt	ctacctcgtg	tacgaatcaa	agcacttaca	tgagggggca	aagtcagaga	1260
cagctgagga	gctgaagaag	gtggctcagg	agctggagga	gaagctaaac	attctcaaca	1320
ataattataa	gattctgcag	gcggaaccaag	aactgtgacc	acagggcagg	gcagccacca	1380
ggagagatat	gcctggcagg	ggccaggaca	aaatgcaaac	tttttttttt	ttctgagaca	1440
gagtcttgct	ctgtcgccaa	gttggagtgc	aatggtgcga	tctcagctca	ctgcaagctc	1500
tgcctcccgt	gttcaagcga	ttctcctgcc	ttggcctccc	aagtagctgg	gactacaggg	1560
gcctaccacc	atgccagct	aattttttgt	tttttaatat	agatgggggt	tcaccatggt	1620
ggccaggatg	gtctcgatct	cctgacctct	tgatctgccc	accttggcct	cccaaagtgc	1680
tgggattaca	ggcgtgagcc	atcgcttttg	acccaaatgc	aaacatttta	ttagggggat	1740
aaagaggggtg	aggtaaagtt	tatggaactg	agtgttaggg	actttggcat	ttccatagct	1800
gagcacagca	ggggaggggt	taatgcagat	ggcagtgcag	caaggagaag	gcaggaacat	1860
tggagcctgc	aataagggaa	aaatgggaac	tggagagtgt	ggggaatggg	aagaagcagt	1920
ttactttaga	ctaaagaata	tattgggggg	ccgggtgtag	tggctcatgc	ctgtaatccg	1980
agcacttttg	gaggccaagg	cgggcggatc	acgaggtcag	gagatcgaga	ccatcctggc	2040
taacacagtg	aaaccccgtc	tctactaaaa	atacaaaaaa	ttagccgggc	atggtggcgg	2100
gcgctttag	ttccagctaa	ctgggcggct	gaggcaggag	aatggcgtag	acctgggagg	2160
tggagcttgc	agttagccga	gatatcgcca	ctgcactcca	gcctgggtga	cagagcgaga	2220
ctccatctca	aaaaaaaaaa	aaaaaagaat	atattgacgg	aagaatagag	aggaggcttg	2280
aaggaaccag	caatgagaag	gccaggaaaa	gaaagagctg	aaaatggaga	aagcccaaga	2340
gttagaacag	ttggatacag	gagaagaaac	agcggctcca	ctacagaccc	agccccaggt	2400
tcaatgtcct	ccgaagaatg	aagtctttcc	ctgggtgatg	tcccctgccc	tgtctttcca	2460
gcattccactc	tcccttgtcc	tcctgggggc	atatctcagt	caggcagcgg	cttcctgatg	2520
atggtcattg	gggtggttgt	catgtgatgg	gtcccctcca	ggttactaaa	gggtgcatgt	2580
cccctgcttg	aacactgaag	ggcagggtgt	gggccatggc	catggtcccc	agctgaggag	2640
cagggtgtccc	tgagaaccca	aacttcccag	agagtatgtg	agaaccaacc	aatgaaaaca	2700
gtcccatcgc	tcttaccggg	taagtaaaca	gtcagaaaat	tagcatgaaa	gcagtttagc	2760
attgggagga	agctcagatc	tctagagctg	tcttgtcgcc	gcccaggatt	gacctgtgtg	2820
taagtcccaa	taaactcacc	tactcatcaa	gctgga			2856

<210> 61

<211> 1655

<212> DNA

<213> NM\_002164.3| Homo sapiens indoleamine-pyrrole 2,3 dioxygenase (INDO), mRNA

```

<400> 61
aatttctcac tgcccctgtg ataaactgtg gtcactggct gtggcagcaa ctattataag      60
atgctctgaa aactcttcag acactgaggg gcaccagagg agcagactac aagaatggca      120
cacgctatgg aaaactcctg gacaatcagt aaagagtacc atattgatga agaagtgggc      180
tttgctctgc caaatccaca ggaaaatcta cctgattttt ataatgactg gatgttcatt      240
gctaaacatc tgcctgatct catagagtct ggccagcttc gagaaagagt tgagaagtta      300
aacatgctca gcattgatca tctcacagac cacaagtcac agcgccctgc acgtctagtt      360
ctgggatgca tcaccatggc atatgtgtgg ggcaaaggctc atggagatgt ccgtaaggctc      420
ttgccaagaa atattgctgt tccttactgc caactctcca agaaactgga actgcctcct      480
attttggttt atgcagactg tgtcttggca aactggaaga aaaaggatcc taataagccc      540
ctgacttatg agaacatgga cgttttgttc tcatttcgtg atggagactg cagtaaagga      600
ttcttcctgg tctctctatt ggtggaaata gcagctgctt ctgcaatcaa agtaattcct      660
actgtattca aggcaatgca aatgcaagaa cgggacactt tgctaaaggc gctgttggaa      720
atagcttctt gcttggagaa agcccttcaa gtgtttcacc aaatccacga tcatgtgaac      780
ccaaaagcat ttttcagtgt tcttcgcata tatttgtctg gctggaaagg caacccccag      840
ctatcagacg gtctgggtgta tgaagggttc tgggaagacc caaaggagtt tgcagggggc      900
agtgagggcc aaagcagcgt ctttcagtgc tttgacgtcc tgctgggcat ccagcagact      960
gctgggtggag gacatgctgc tcagttcctc caggacatga gaagatatat gccaccagct     1020
cacaggaact tcctgtgctc attagagtca aatccctcag tccgtgagtt tgcctttca     1080
aaagggtgatg ctggcctgcg ggaagcttat gacgcctgtg tgaaagctct ggtctccctg     1140
aggagctacc atctgcaaat cgtgactaag tacatcctga ttcttgcaag ccagcagcca     1200
aaggagaata agacctctga agacccttca aaactggaag ccaaaggaac tggaggcact     1260
gatttaatga atttcctgaa gactgtaaga agtacaactg agaaatccct tttgaaggaa     1320
ggttaatgta acccaacaag agcacatttt atcatagcag agacatctgt atgcattcct     1380
gtcattaccc attgtaacag agccacaaac taatactatg caatgtttta ccaataatgc     1440
aatacaaaag acctcaaaat acctgtgcat ttctttagg aaaacaacaa aaggtaatta     1500
tgtgtaatta tactagaagt tttgtaatct gtatcttattc attggaataa aatgacattc     1560
aataaataaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa     1620
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa

```

<210> 62

<211> 2242

<212> DNA

<213> NM\_021784.3| Homo sapiens forkhead box A2 (FOXA2), transcript variant 1, mRNA

<400> 62

cccgccttact	tccaactacc	gcctccggcc	tgcccagggga	gagagagggga	gtggagccca	60
gggagagggga	gcgcgagaga	gggagggagg	aggggacggt	gctttggctg	actttttttt	120
aaaagaggggt	gggggtgggg	ggtgattgct	ggtcgtttgt	tgtggctggt	aaattttaaa	180
ctgccatgca	ctcggcttcc	agtatgctgg	gagcggtgaa	gatggaaggg	cacgagccgt	240
ccgactggag	cagctactat	gcagagcccc	agggctactc	ctccgtgagc	aacatgaacg	300
ccggcctggg	gatgaacggc	atgaacacgt	acatgagcat	gtcggcggcc	gccatgggca	360
gcggctcggg	caacatgagc	gcgggctcca	tgaacatgtc	gtcgtacgtg	ggcgctggca	420
tgagcccgtc	cctggcgggg	atgtcccccg	gcgcggggcg	catggcgggc	atgggcgggt	480
cggccggggc	ggccggcgtg	gcgggcatgg	ggccgcactt	gagtcccagc	ctgagcccgc	540
tcggggggca	ggcggccggg	gccatgggcg	gcctggcccc	ctacgccaac	atgaactcca	600
tgagccccat	gtacgggagc	gcgggcctga	gccgcgcccc	cgaccccaag	acctacaggc	660
gcagctacac	gcacgcaaag	ccgccctact	cgtacatctc	gctcatcacc	atggccatcc	720
agcagagccc	caacaagatg	ctgacgctga	gcgagatcta	ccagtggatc	atggacctct	780
tcccccttcta	ccggcagaac	cagcagcgct	ggcagaactc	catccgccac	tcgctctcct	840
tcaacgactg	tttctgaag	gtgccccgct	cgcccgacaa	gcccggcaag	ggctccttct	900
ggaccctgca	ccctgactcg	ggcaacatgt	tcgagaacgg	ctgctacctg	cgccgccaga	960
agcgcttcaa	gtgcgagaag	cagctggcgc	tgaaggaggc	cgaggcgccc	gccggcagcg	1020
gcaagaaggc	ggccgcccga	gcccaggcct	cacaggctca	actcggggag	gccgcccggc	1080
cggcctccga	gactccggcg	ggcaccgagt	cgcctcactc	gagcgccctc	ccgtgccagg	1140
agcacaagcg	agggggcctg	ggagagctga	aggggacgcc	ggctgcggcg	ctgagcccc	1200
cagagccggc	gcccctctcc	gggcagcagc	agcaggccgc	ggccccacctg	ctgggcccgc	1260
cccaccaccc	gggcctgccc	cctgaggccc	acctgaagcc	ggaacaccac	tacgccttca	1320
accacccggt	ctccatcaac	aacctcatgt	cctcggagca	gcagcaccac	cacagccacc	1380
accaccacca	acccacaaaa	atggacctca	aggcctacga	acagggtgatg	cactaccccg	1440
gctacgggtc	ccccatgcct	ggcagcttgg	ccatggggccc	ggtcacgaac	aaaacggggc	1500
tggacgcctc	gcccctggcc	gcagatacct	cctactacca	gggggtgtac	tcccgggcca	1560
ttatgaactc	ctcttaagaa	gacgacggct	tcaggcccgg	ctaactctgg	caccccggt	1620
cgaggacaag	tgagagagca	agtgggggtc	gagactttgg	ggagacgggtg	ttgcagagac	1680
gcaagggaga	agaaatccat	aacaccccc	ccccaacacc	cccaagacag	cagtcttctt	1740
cacccgctgc	agccgttccg	tcccaaacag	agggccacac	agatacccca	cgttctatat	1800



<210> 64

<211> 5243

<212> DNA

<213> NM\_001165.3| Homo sapiens baculoviral IAP repeat-containing 3 (BIRC3), transcript variant 1, mRNA

<400> 64

```
agcgtgagac tcgcgccctc cggcacggaa aaggccaggc gacaggtgtc gcttgaaaag    60
actgggcttg tccttgctgg tgcattgcgtc gtcggcctct gggcagcagg ttacaaaagg    120
aggaaaacga cttcttctag attttttttt cagtttcttc tataaatcaa aacatctcaa    180
aatggagacc taaaatcctt aaagggactt agtctaattc cgggaggtag ttttgtgcat    240
gggtaaaciaa attaatgatt aactgggtgt ttactatcca aagaatgcta attttataaa    300
catgatcgag ttatataagg tataccataa tgagtttgat tttgaatttg atttgtggaa    360
ataaaggaaa agtgattcta gctggggcat attgttaaag cttttttttc agagttggcc    420
aggcagtctc ctactggcac attctcccat tatgtagaat agaaatagta cctgtgtttg    480
ggaaagattt taaaatgagt gacagttatt tggaacaaag agctaataat caatccactg    540
caaattaaag aaacatgcag atgaaagttt tgacacatta aaatacttct acagtgacaa    600
agaaaaatca agaacaaagc tttttgatat gtgcaacaaa ttttagaggaa gtaaaaagat    660
aaatgtgatg attgggtcaag aaattatcca gttattttaca aggccactga tatttttaaac    720
gtccaaaagt ttgttttaaat gggctgttac cgctgagaat gatgaggatg agaatgatgg    780
ttgaaggtta ctttttagga aatgaagaaa cttagaaaat taatataaag acagtgatga    840
atacaaagaa gattttttata acaatgtgta aaatttttgg ccagggaag gaatattgaa    900
gttagatata attacttacc tttgagggaa ataattgttg gtaatgagat gtgatgtttc    960
tcctgccacc tggaacaaaa gcattgaagt ctgcagttga aaagcccaac gtctgtgaga   1020
tccaggaaac catgcttgca aaccactggg aaaaaaaaaa aaaaaaaagcca   1080
cagtgacttg cttattgggtc attgctagta ttatcgactc agaacctctt tactaatggc   1140
tagtaaatca taattgagaa attctgaatt ttgacaaggc ctctgctgtt gaaatggtaa   1200
atttattatt ttttttgtca tgataaattc tggttcaagg tatgctatcc atgaaataat   1260
ttctgaccaa aactaaattg atgcaatttg attatccatc ttagcctaca gatggcatct   1320
ggtaactttt gactgtttta aaaaataaat ccactatcag agtagatttg atgttggtct   1380
cagaaacatt tagaaaaaca aaagttcaaa aatgttttca ggaggtgata agttgaataa   1440
ctctacaatg ttagttcttt gaggggggaca aaaaatttaa aatctttgaa aggtcttatt   1500
ttacagccat atctaaatta tcttaagaaa atttttaaca aagggaatga aatataatc   1560
atgattctgt ttttccaaaa gtaacctgaa tatagcaatg aagttcagtt ttgttattgg   1620
tagtttgggc agagtctctt tttgcagcac ctgttgtcta ccataattac agaggacatt   1680
```



tccatgttct agccaagtat actattagaa taaaaaaact taacattgag ttgcttcaac	1740
agcatgaaac tgagtccaaa agaccaaatg aacaaacaca ttaatctctg attattttatt	1800
ttaaatagaa tattttaattg tgtaagatct aatagtatca ttatacttaa gcaatcatat	1860
tcctgatgat ctatgggaaa taactattat ttaattaata ttgaaaccag gttttaagat	1920
gtgttagcca gtcctgttac tagtaaactc ctttatttgg agagaaattt tagattgttt	1980
tgttctcctt attagaagga ttgtagaaag aaaaaaatga ctaattggag aaaaattggg	2040
gatatatcat atttcactga attcaaaatg tcttcagttg taaatcttac cattatttta	2100
cgtacctcta agaaataaaa gtgcttctaa ttaaaatatg atgtcattaa ttatgaaata	2160
cttcttgata acagaagttt taaaatagcc atcttagaat cagtgaataa tggtaatgta	2220
ttattttcct cttttgagtt aggtcttgtg ctttttttc ctggccacta aatttcacaa	2280
tttccaaaaa gcaaaataaa catattctga atatttttgc tgtgaaacac ttgacagcag	2340
agctttccac catgaaaaga agcttcatga gtcacacatt acatctttgg gttgattgaa	2400
tgccactgaa acattctagt agcctggaga agttgacctt cctgtggaga tgcctgccat	2460
taaatggcat cctgatggct taatacacat cactcttctg tgaagggttt taattttcaa	2520
cacagcttac tctgtagcat catgtttaca ttgtatgtat aaagattata caaagggtgca	2580
attgtgtatt tcttccttaa aatgtatcag tataggattt agaatctcca tgttgaaact	2640
ctaaatgcat agaaataaaa ataataaaaa atttttcatt ttggcttttc agcctagtat	2700
taaaactgat aaaagcaaag ccatgcacaa aactacctcc ctagagaaag gctagtcctt	2760
tttcttcccc attcatttca ttatgaacat agtagaaaac agcatattct tatcaaattt	2820
gatgaaaagc gccaacacgt ttgaactgaa atacgacttg tcatgtgaac tgtaccgaat	2880
gtctacgtat tccacttttc ctgctggggt tctgtctca gaaaggagtc ttgctcgtgc	2940
tggtttctat tacactgggtg tgaatgacaa ggtcaaatgc ttctgttgtg gcctgatgct	3000
ggataactgg aaaagaggag acagtcctac tgaaaagcat aaaaagtgtg atcctagctg	3060
cagattcggt cagagtctaa attccgttaa caacttgga gctacctctc agcctacttt	3120
tccttcttca gtaacaaatt ccacacactc attacttccg ggtacagaaa acagtggata	3180
tttccgtggc tcttattcaa actctccatc aaatcctgta aactccagag caaatcaaga	3240
tttttctgcc ttgatgagaa gttcctacca ctgtgcaatg aataacgaaa atgccagatt	3300
acttactttt cagacatggc cattgacttt tctgtcgcca acagatctgg caaaagcagg	3360
cttttactac ataggacctg gagacagagt ggcttgcttt gcctgtgggtg gaaaattgag	3420
caattgggaa ccgaaggata atgctatgtc agaacacctg agacattttc ccaaagccc	3480
atztatagaa aatcagcttc aagacacttc aagatacaca gtttctaatac tgagcatgca	3540
gacacatgca gcccgcttta aaacattctt taactggccc tctagtgttc tagttaatcc	3600
tgagcagctt gcaagtgcgg gtttttatta tgtgggtaac agtgatgatg tcaaagtctt	3660

ttgctgtgat ggtggactca ggtggtggga atctggagat gatccatggg ttcaacatgc	3720
caagtggttt ccaagggtgtg agtacttgat aagaattaaa ggacaggagt tcatccgtca	3780
agttcaagcc agttaccctc atctacttga acagctgcta tccacatcag acagcccagg	3840
agatgaaaat gcagagtcac caattatcca ttttgaacct ggagaagacc attcagaaga	3900
tgcaatcatg atgaatactc ctgtgattaa tgctgccgtg gaaatgggct ttagtagaag	3960
cctggtaaaa cagacagttc agagaaaaat cctagcaact ggagagaatt atagactagt	4020
caatgatctt gtgttagact tactcaatgc agaagatgaa ataagggag aggagagaga	4080
aagagcaact gaggaaaaag aatcaaatga tttattatta atccggaaga atagaatggc	4140
actttttcaa catttgactt gtgtaattcc aatcctggat agtctactaa ctgccggaat	4200
tattaatgaa caagaacatg atgttattaa acagaagaca cagacgtctt tacaagcaag	4260
agaactgatt gatacgattt tagtaaaagg aaatattgca gccactgtat tcagaaactc	4320
tctgcaagaa gctgaagctg tggttatatga gcattttatt gtgcaacagg acataaaata	4380
tattcccaca gaagatgttt cagatctacc agtggaagaa caattgcgga gactacaaga	4440
agaaagaaca tgtaaagtgt gtatggacaa agaagtgtcc atagtgttta ttccttgtgg	4500
tcacttagta gtatgcaaag attgtgctcc ttccttaaga aagtgtccta tttgtaggag	4560
tacaatcaag ggtacagttc gtacatttct ttcattgaaga agaaccacaaa catcgtctaa	4620
actttagaat taatttatta aatgtattat aactttaact tttatcctaa tttggtttcc	4680
ttaaaatttt tattttatta caactcaaaa aacattgttt tgtgtaacat atttatatat	4740
gtatctaaac catatgaaca tatatttttt agaaactaag agaattgatag gcttttggtc	4800
ttatgaacga aaaagaggta gcactacaaa cacaatattc aatcaaaatt tcagcattat	4860
tgaaattgta agtgaagtaa aacttaagat atttgagtta acctttaaga attttaaata	4920
ttttggcatt gtactaatac cggaacatg aagccagggtg tgggtggtatg tgcctgtagt	4980
cccaggctga ggcaagagaa ttacttgagc ccaggagttt gaatccatcc tgggcagcat	5040
actgagaccc tgccttttaa aacaaacaga acaaaaacaa aacaccaggg acacatttct	5100
ctgtcttttt tgatcagtg cctatacatc gaagggtgac atatatgttg aatgacattt	5160
tagggacatg gtgtttttat aaagaattct gtgagaaaaa atttaataaa gcaacaaaaa	5220
ttactcttaa aaaaaaaaaa aaa	5243

<210> 65

<211> 3850

<212> DNA

<213> NM\_005682.4| Homo sapiens G protein-coupled receptor 56 (GPR56), transcript variant 1, mRNA

<400> 65

agactgggtg cctgtggccc tgggaggagg tgggaagggga ggagcaggcc acacaggcac	60
aggccggtga gggacctgcc cagacctgga gggctctcgt ctgtcacaca ggctggagt	120
cagtgggtg atcttggctc atcgtaacct ccacctcccg ggttcaagtg attctcatgc	180
ctcagcctcc cgagtagctg ggattacagg tggtgacttc caagagtga tccgtcggag	240
gaaaatgact cccagtcgc tgctgcagac gacactgttc ctgctgagtc tgctcttct	300
ggtccaaggt gcccacggca ggggccacag ggaagacttt cgcttctgca gccagcggaa	360
ccagacacac aggagcagcc tccactacaa acccacacca gacctgcgca tctccatcga	420
gaactccgaa gaggccctca cagtccatgc ccctttccct gcagcccacc ctgcttccc	480
atccttccct gaccccaggg gcctctacca cttctgcctc tactggaacc gacatgctgg	540
gagattacat cttctctatg gcaagcgtga cttcttgctg agtgacaaag cctctagcct	600
cctctgcttc cagcaccagg aggagagcct ggctcagggc ccccgctgt tagccacttc	660
tgctacctcc tgggtggagcc ctcagaacat cagcctgccc agtgccgcca gcttcacctt	720
ctccttccac agtcctcccc acacggccgc tcacaatgcc tcggtggaca tgtgcgagct	780
caaaagggac ctccagctgc tcagccagtt cctgaagcat cccagaagg cctcaaggag	840
gccctcggct gcccccgcca gccagcagtt gcagagcctg gagtcgaaac tgacctctgt	900
gagattcatg ggggacatgg tgtccttcga ggaggaccgg atcaacgcca cgggtgtggaa	960
gctccagccc acagccggcc tccaggacct gcacatccac tcccggcagg aggaggagca	1020
gagcgagatc atggagtact cgggtgctgt gcctcgaaca ctcttccaga ggacgaaagg	1080
ccggagcggg gaggctgaga agagactcct cctgggtggac ttcagcagcc aagccctgtt	1140
ccaggacaag aattccagcc aagtcctggg tgagaagggtc ttggggattg tggtagagaa	1200
caccaaagta gccaacctca cggagcccgt ggtgtcact ttcagcacc agctacagcc	1260
gaagaatgtg actctgcaat gtgtgttctg ggttgaagac cccacattga gcagcccggg	1320
gcattggagc agtgctgggt gtgagaccgt caggagagaa acccaaacat cctgcttctg	1380
caaccacttg acctactttg cagtgtgat ggtctcctcg gtggaggtgg acgccgtgca	1440
caagcactac ctgagcctcc tctcctacgt gggctgtgtc gtctctgccc tggcctgcct	1500
tgctaccatt gccgcctacc tctgtctcag ggtgcccctg ccgtgcagga ggaaacctcg	1560
ggactacacc atcaaggtgc acatgaacct gctgtggcc gtcttctgc tggacacgag	1620
cttctgtctc agcgagccgg tggccctgac aggctctgag gctggctgcc gagccagtgc	1680
catcttctct cacttctccc tgctcacctg cctttcctgg atgggcctcg aggggtacaa	1740
cctctaccga ctggtgggtg aggtctttgg cacctatgtc cctggctacc tactcaagct	1800
gagcgccatg ggctggggct tccccatctt tctggtgacg ctggtggccc tgggtgatgt	1860
ggacaactat ggccccatca tcttggtgtg gcataggact ccagagggcg tcctctaccc	1920
ttccatgtgc tggatccggg actccctggg cagctacatc accaacctgg gcctcttcag	1980
cctggtgttt ctgttcaaca tggccatgct agccaccatg gtggtgcaga tcctgcggct	2040

gcgccccac acccaaaagt ggtcacatgt gctgacactg ctgggcctca gcctggctct	2100
tgccctgccc tgggccttga tcttcttctc ctttgcttct ggcaccttcc agcttgctgt	2160
cctctacctt ttcagcatca tcacctcctt ccaaggcttc ctcatcttca tctgggtactg	2220
gtccatgcgg ctgcaggccc ggggtggccc ctcccctctg aagagcaact cagacagcgc	2280
caggctcccc atcagctcgg gcagcacctc gtccagccgc atctaggcct ccagcccacc	2340
tgcccatgtg atgaagcaga gattcggcct cgtcgcacac tgcctgtggc ccccgagccc	2400
ggcccagccc caggccagtc agccgcagac tttggaaagc ccaacgacca tggagagatg	2460
ggccgttgcc atggtggacg gactccccgg ctgggctttt gaattggcct tggggactac	2520
tcggctctca ctcagctccc acgggactca gaagtgcgc gccatgctgc ctagggtact	2580
gtccccacat ctgtcccaac ccagctggag gcctggcttc tccttacaac ccctgggccc	2640
agccctcatt gctgggggccc aggccttggg tcttgagggt ctggcacatc cttaatcctg	2700
tgccctgcc tgggacagaa atgtggctcc agttgctctg tctctcgtgg tcaccctgag	2760
ggcactctgc atcctctgtc attttaacct cagggtggac ccagggcgaa tggggcccag	2820
ggcagacctt caggggcaga gccctggcgg aggagaggcc ctttgccagg agcacagcag	2880
cagctcgcct acctctgagc ccaggcccc tccttcctc agccccccag tcctccctcc	2940
atcttccctg gggttctct cctctcccag ggctctcttg ctcttctgtt cacagctggg	3000
ggtccccgat tccaatgctg ttttttgggg agtggtttcc aggagctgcc tgggtgtctgc	3060
tgtaaatgtt tgtctactgc acaagcctcg gcctgcccct gagccaggct cggtagccgat	3120
gcgtgggctg ggctaggctc ctctgtccat ctgggccttt gtatgagctg cattgccctt	3180
gctcaccctg accaagcaca cgcctcagag gggccctcag cctctcctga agccctcttg	3240
tggcaagaac tgtggaccat gccagtcccg tctggtttcc atcccaccac tccaaggact	3300
gagactgacc tcctctgggtg aacttgccct agggcctgac actctcctaa gaggttctct	3360
ccaagcccc aaatagctcc aggcgcctc ggccgcccac catggttaat tctgtccaac	3420
aaacacacac gggtagattg ctggcctgtt gtaggtggta gggacacaga tgaccgacct	3480
ggtcactcct cctgccaaca ttcagtctgg tatgtgaggc gtgcgtgaag caagaactcc	3540
tggagctaca gggacagggg gccatcattc ctgcctggga atcctggaag acttcttgca	3600
ggagtcagcg ttcaatcttg accttgaaga tgggaaggat gttcttttta cgtaccaatt	3660
cttttgtctt ttgatattaa aaagaagtac atgttcattg tagagaattt ggaaactgta	3720
gaagagaatc aagaagaaaa ataaaaatca gctgttgtaa tcacctagca aactggaaaa	3780
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	3840
aaaaaaaaa	3850

<210> 66

<211> 372

<212> DNA

<213> NM\_005953.2| Homo sapiens metallothionein 2A (MT2A), mRNA

<400> 66  
agtcccagcg aacccgcgtg caacctgtcc cgactctagc cgcctcttca gcacgccatg 60  
gatcccaact gtcctgctgc cgccggtgac tcctgcacct gcgccgggttc ctgcaaattg 120  
aaagagtgc aatgcacttc gtgcaagaaa agctgctgct cctgctgccc tgtgggctgt 180  
gccaagtgtg cccaaggctg catctgcaaa ggggcgtcgg acaagtgcag ctgctgcgcc 240  
tgatgctggg acagccccgc tcccagatgt aaagaacgcg acttccacaa acctggattt 300  
tttatgtaca accctgaccg tgaccgtttg ctatattcct ttttctatga aataatgtga 360  
atgataataa aa 372

<210> 67

<211> 4180

<212> DNA

<213> NM\_015002.1| Homo sapiens F-box protein 21 (FBXO21), transcript variant 2, mRNA

<400> 67  
gtacgcggac aagatggcgg cggcagcagt cgacagcgcg atggaggtgg tgccggcgct 60  
ggcggaggag gccgcgccgg aggtagcggg cctcagctgc ctctgcaacc tgccgggtga 120  
ggtgctggag tacatcctgt gctgcggctc gctgacggcc gccgacatcg gccgtgtctc 180  
cagcacctgc cggcggctgc gcgagctgtg ccagagcagc gggaagggtg ggaaggagca 240  
gttccgggtg aggtggcctt cccttatgaa aactacagc cccaccgact acgtcaattg 300  
gttggaagag tataaagttc ggcaaaaagc tgggttagaa gcgcggaaga ttgtagcctc 360  
gttctcaaag aggttctttt cagagcacgt tccttgtaat ggcttcagtg acattgagaa 420  
ccttgaagga ccagagattt tttttgagga tgaactgggtg tgtatcctaa atatggaagg 480  
aagaaaagct ttgacctgga aatactacgc aaaaaaatt ctttactacc tgcggcaaca 540  
gaagatctta aataatctta aggcctttct tcagcagcca gatgactatg agtcgtatct 600  
tgaagggtgct gtatatattg accagtactg caatcctctc tccgacatca gcctcaaaga 660  
catccaggcc caaattgaca gcatcgtgga gcttgtttgc aaaacccttc ggggcataaa 720  
cagtcgccac cccagcttgg ccttcaaggc aggtgaatca tccatgataa tggaaataga 780  
actccagagc cagggtgctgg atgccatgaa ctatgtcctt tacgaccaac tgaagttcaa 840  
ggggaatcga atggattact ataatgccct caacttatat atgcatcagg ttttgattcg 900  
cagaacagga atcccaatca gcatgtctct gctctatttg acaattgctc ggcagttggg 960

agtcccaetg	gagcctgtca	acttcccaag	tcactttctta	ttaaggtggt	gccaaaggcgc	1020
agaaggggcg	accctggaca	tctttgacta	catctacata	gatgcttttg	ggaaaggcaa	1080
gcagctgaca	gtgaaagaat	gcgagtactt	gatcggccag	cacgtgactg	cagcactgta	1140
tggggtggtc	aatgtcaaga	aggtgttaca	gagaatggtg	ggaaacctgt	taagcctggg	1200
gaagcgggaa	ggcatcgacc	agtcatacca	gctcctgaga	gactcgctgg	atctctatct	1260
ggcaatgtac	ccggaccagg	tgcagcttct	cctcctccaa	gccaggcttt	acttccacct	1320
gggaatctgg	ccagagaagg	tgcttgacat	cctccagcac	atccaaaccc	tagacccggg	1380
gcagcacggg	gcggtgggct	acctggtgca	gcacactcta	gagcacattg	agcgcaaaaa	1440
ggaggaggtg	ggcgtagagg	tgaagctgcg	ctccgatgag	aagcacagag	atgtctgcta	1500
ctccatcggg	ctcattatga	agcataagag	gtatggctat	aactgtgtga	tctacggctg	1560
ggacccacc	tgcattgatg	gacacgagtg	gatccggaac	atgaacgtcc	acagcctgcc	1620
gcacggccac	caccagcctt	tctataacgt	gctggtggag	gacggctcct	gtcgatacgc	1680
agcccaagaa	aacttggaa	ataacgtgga	gcctcaagaa	atctcacacc	ctgacgtggg	1740
acgctatttc	tcagagttta	ctggcactca	ctacatccca	aacgcagagc	tggagatccg	1800
gtatccagaa	gatctggagt	ttgtctatga	aacggtgcag	aatatttaca	gtgcaaagaa	1860
agagaacata	gatgagtaaa	gtctagagag	gacattgcac	ctttgctgct	gctgctatct	1920
tccaagagaa	cgggactccg	gaagaagacg	tctccacgga	gccctcggga	cctgctgcac	1980
caggaaagcc	actccaccag	tagtgctggt	tgccctctac	taagtttaaa	taccgtgtgc	2040
tcttccccag	ctgcaaagac	aatgttgctc	tccgcctaca	ctagtgaatt	aatctgaaa	2100
gcactgtgtc	agtggcatgg	cttgatgct	tgtcctgtgg	tgacagtttg	tgacattctg	2160
tcttcatgag	gtctcacagt	cgacgctcct	gtaatcattc	tttgatttca	ctccattccc	2220
ctgtctgtct	gcatttgtct	cagaacattt	ccttggtggt	acagatgggg	ttatgcattt	2280
gcaataattt	ccttctgatt	tctctgtgga	acgtgttcgg	tcccagtgta	ggactgtgtg	2340
tctttttacc	ctgaagttag	ttgcatattc	agaggtaaa	ttgtgtgcta	tcttggcagc	2400
atcttagaga	tggagacatt	aacaagctaa	tggttaattag	aatcatttga	atttattttt	2460
ttctaatatg	tgaaacacag	atttcaagtg	ttttatcttt	tttttttaaa	tttaaatggg	2520
aatataacac	agttttccct	tccatattcc	tctcttgagt	ttatgcacat	ctctataaat	2580
cattagtttt	ctattttatt	acataaaatt	cttttagaaa	atgcaaata	tgaaactttgt	2640
gaatggattt	ttccatactc	atctacaatt	cctccatttt	aaatgactac	ttttattttt	2700
taatttaaaa	aatctacttc	agtatcatga	gtaggtctta	catcagtgat	gggttctttt	2760
tgtagtgaga	catacaaata	tgatgttaat	gtttgctctt	agaagtcata	ctccatgggtc	2820
ttcaaagacc	aaaaaatgag	gttttgcctt	tgtaatcagg	aaaaaaaaaa	attaatgaac	2880
cttaaaaaaa	aaaaaaaaag	ttttgaaggg	aaaaaaagt	gtttcacacc	tcttgttatt	2940

ccttagagtc	acttcaaggc	ctgtttgaat	gtggcagggt	agaaagagag	agaatgtctt	3000
tcatttgaag	agtgttgac	ttgtgtgaaa	ggagatgtgc	gtgttggaat	ctgcttttcc	3060
aagccgccag	ggcctgacg	gcagcaggac	gaagcctgtt	gtggcgtctt	ctgggaaagc	3120
ctgaccgtgt	gttcggacgg	cactggctcc	tttccgaagt	tctcagtaac	tgagcccaga	3180
gtaactgcac	gcctttgtgc	agctctggag	ctccaccaac	tctcggcctg	ccagttctca	3240
agcgagctaa	tcttgtcatt	aatcgataga	agctaacttc	cgaagttagg	acctagttac	3300
tttgctctca	acatttaaaa	taatgcagtt	gctctagtga	atggggcggt	aggggcctgt	3360
ctctgcacct	gtctgtccat	ctgcatgcag	tattctcacc	catgttgaat	gcctgctgct	3420
tgtttaccct	ttggaaaccc	tggggtgacc	aagggttgga	aagccacctg	agaccacttc	3480
atagcaaggg	aaggctttta	gcagttacta	gaaagagatg	gggatttggc	ccctggctcc	3540
tccagcctga	atgagctatt	taatccactg	tccatgttcc	tcatcagtca	aatccaaagt	3600
caaaggattt	gaacctgcat	ctggaaacgt	aaccactcac	agcacctggc	ccgccaaggt	3660
tgggaggatt	gtacactact	ttcattttaa	ggggaaagtt	tgataatacg	gaattaatta	3720
atatgaatga	gatgcattaa	taagaacctg	agcatgctga	gagttgcaat	tgttgggttt	3780
ctggtttgat	tgatttcctt	ttttcttaga	cacatcaaag	tcaagaaaga	tggttttacc	3840
tttactgacc	cagctgtaca	tatgtatcta	gactgttttt	aaatgtcttt	cttcatgaat	3900
gcttcatggg	gctccaggaa	gcctgtatca	cctgtgtaag	ttggattttg	ggcactttat	3960
atTTTTtctaa	aaacgtgttt	tggatcctgt	actctaataa	atcataagtt	tcttttttaa	4020
aatTTTtcaa	aactTTTctc	cattTTtaaa	agccctgtta	taaacgttga	actttcacaa	4080
tgttaaaatg	ttaaataattt	ggatatagca	acttcttttc	tcttcaaattg	aatgccaaga	4140
TTTTTTtgta	caatgattaa	taaatggaac	ttatccagag			4180

<210> 68

<211> 6276

<212> DNA

<213> NM\_012156.2| Homo sapiens erythrocyte membrane protein band 4.1-like 1 (EPB41L1), transcript variant 1, mRNA

<400> 68						
agtcggcatc	catcagcggg	cgggggtgtc	gccgaacagg	ctgctccgca	gagcccgcgc	60
cgaccccgcg	ccgccccgcc	ccgcggcctg	cctgccagag	gagccgaggg	ggccgcccct	120
cgcccaacct	gcccgacatg	gggaaccccg	ggcccaggcg	tgctgggtcac	catgacaaca	180
gagacaggcc	ccgactctga	ggtagaagaa	gctcaggagg	aggccccgca	gcagcccagag	240
gctgctgccg	ctgtgaccac	ccctgtgacc	cctgcaggcc	acggccaccc	agaggccaac	300
tccaatgaga	agcatccatc	ccagcaggac	acgcggcctg	ctgaacagag	cctagacatg	360

gaggagaagg	actacagtga	ggccgatggc	ctttcggaga	ggaccacgcc	cagcaaggcc	420
cagaaatcgc	cccagaagat	tgccaagaaa	tacaagagtg	ccatctgccg	ggtcactctg	480
cttgatgcct	cggagtatga	gtgtgaggtg	gagaaacatg	gccggggcca	ggtgctgttt	540
gacctggtct	gtgaacacct	caacctccta	gagaaggact	acttcggcct	gaccttctgt	600
gatgctgaca	gccagaagaa	ctggctggac	ccctccaagg	agatcaagaa	gcagatccgg	660
agtagcccc	ggaattttgc	cttcacagtc	aagttctacc	cgctgatcc	tgcccagctg	720
acagaagaca	tcacaagata	ctacctgtgc	ctgcagctgc	gggcagacat	catcacgggc	780
cggctgccat	gctcctttgt	cacgcatgcc	ctactgggct	cctacgctgt	gcaggctgag	840
ctgggtgact	atgatgctga	ggagcatgtg	ggcaactatg	tcagcgagct	ccgcttcgcc	900
cctaaccaga	cccgggagct	ggaggagagg	atcatggagc	tgcataagac	atataggggg	960
atgacccccg	gagaagcaga	aatccacttc	ttagagaatg	ccaagaagct	ttccatgtac	1020
ggagtagacc	tgcaccatgc	caaggactct	gagggcatcg	acatcatgtt	aggcgtttgt	1080
gccaatggcc	tgctcatcta	ccgggaccgg	ctgagaatca	accgctttgc	ctggcccaag	1140
atcctcaaga	tctcctacaa	gaggagtaac	ttctatatca	agatccggcc	tggggagtat	1200
gagcaatttg	agagcacaat	tggctttaag	ctcccaaacc	accggtcagc	caagagactg	1260
tggaaggctt	gcatcgagca	tcatacatc	ttccggctgg	tgtcccctga	gccccaccc	1320
aagggcttcc	tggatgatgg	ctccaagttc	cggtacagtg	ggaggacca	ggcacagact	1380
cgccaggcca	gcgccctcat	tgaccggcct	gcacccttct	ttgagcgctc	ttccagcaaa	1440
cggtagacca	tgtcccgag	ccttgatgga	gcagagttct	cccggccagc	ctcggtcagc	1500
gagaaccatg	atgcagggcc	tgacgggtgac	aagcgggatg	aggatggcga	gtctgggggg	1560
caacggtcag	aggctgagga	gggagaggtc	aggactccaa	ccaagatcaa	ggagctaaag	1620
ccggagcagg	aaaccacgcc	gagacacaag	caggagtctt	tagacaagcc	agaagatgtc	1680
ttgctgaagc	accaggccag	catcaatgag	ctcaaaagga	ccctgaagga	gccaacagc	1740
aaactcatcc	accgggatcg	agactgggaa	cgggagcgca	ggctgccctc	ctccccgcc	1800
tccccctccc	ccaagggcac	ccctgagaaa	gccaatgaga	gagcagggct	gagggagggc	1860
tccgaggaga	aagtcaaacc	accacgtccc	cggg'ccccag	agagtgcac	aggcgatgag	1920
gaccaggacc	aggagagggg	cacggtgttc	ctgaaggaca	accacctggc	cattgagcgc	1980
aagtgtcca	gcatacagg	cagctctacg	tctagcctgg	aggctgaggt	ggacttcacg	2040
gtcattggtg	actaccatgg	cagcgccttc	gaagacttct	cccgcagcct	gcctgagctc	2100
gaccgggaca	aaagcgactc	ggacactgag	ggcctgctgt	tctcccggga	tctcaacaag	2160
ggggccccc	gccaggatga	tgagtctggg	ggcattgagg	acagcccggg	tcgagggggc	2220
tgctccaccc	cggatatgcc	ccagtttgag	cccgtgaaaa	cagaaaccat	gactgtcagc	2280
agtctggcca	ttagaaagaa	gattgagccg	gaggccgtac	tgagaccag	agtctccgct	2340
atggataaca	cccagcaggt	tgatgggagt	gcctcagtgg	ggagggagtt	catagcaacc	2400



actccctcca	tcaccacgga	gaccatatcg	accaccatgg	agaacagtct	caagtccggg	2460
aagggggcag	ctgccatgat	cccaggccca	cagacggtgg	ccacggaaat	ccgttctctt	2520
tctccgatca	tcgggaaaga	tgtcctcacc	agcacctacg	gcgccactgc	ggaaaccctc	2580
tcaacctcca	ccaccaccca	tgtcaccaaa	actgtgaaag	gaggggtttc	tgagacaagg	2640
atcgagaagc	gaatcatcat	tactggggat	gaagatgtcg	atcaagacca	ggccctggct	2700
ttggccatca	aggaggccaa	actgcagcat	cctgatatgc	tggtaaccaa	agctgtcgta	2760
tacagagaaa	cagacccatc	cccagaggag	agggacaaga	agccacagga	atcctgacct	2820
ctgtgaagag	atcctggcat	ttctggtcca	acccaagcca	gagaaccatt	aagaaggggc	2880
cttcattctg	gatttctcga	cgcaacactg	acgtcccagc	tgcgacgtac	tgtcactgat	2940
gagagactgg	gaagggaaaa	gcatatatat	atagatatat	agagatatag	atatatatac	3000
aggaaacacc	gcattccttg	actgctgctg	gggctggcag	agcagttggc	tgacagcaac	3060
aaccgacatc	tgaacaccta	catttccttt	gcagacaaat	tgaagaactg	gtgggatttt	3120
tttcaagaaa	aaaaattata	taataactat	aatcccttgc	tcaccccttt	ccccgccaa	3180
ataagaaacg	caagccagac	cacgatgatt	gtagaagtcc	ctcccgccct	ggttctgcac	3240
gttacagtta	gcagacgagc	aattccattt	gttcttctcc	agcatctcta	aggccactt	3300
gaatgcaaag	gaaaacactt	gcacagcaaa	gcaagagaag	tcacagcagc	aagacacgca	3360
cagtcaacca	ttttccgaga	aaaaaagaaa	attccccact	tggaaagaaa	gaggaggaaac	3420
actggattct	tactttctgg	atcttgacac	tgggctgcaa	aacctacctt	cctctctccc	3480
gcctcccttc	accctcaact	ctcaatgtct	tgtgtgcatt	ttctgtctcg	gctccctcct	3540
cccccttccc	ccttccccca	ccccacaccc	ttcacctctt	gtgtcctggg	ccttctgagg	3600
gccactgcag	atgactctcc	tttgaaatga	gaaaaagaaa	agaaagcaag	aacagaaaac	3660
gaagccacag	gaaggggaag	agacattgta	tgcttatggg	ttctcattat	gaagggtgcag	3720
cttgtaggag	gtttgtacgg	atgtgctttg	aagttatgta	tattacatat	aacaggaaaa	3780
aatattaaaa	taaacagtgc	tggttaagtat	gaagctgaca	ttctaaaatt	ataattatct	3840
gactgtgatt	gatgtatcct	gaggttccta	gatctcactg	aactggccca	gctaaggaga	3900
cctggactct	gggtgtgggt	tggctcacag	taggggctga	cgggttcagt	gtagtaatac	3960
tgtgtgtggg	gtttgtaatt	ggttgattgg	tggggagggg	tggggggccc	taatggagag	4020
gtgtgggttt	ggcaagaaag	aagcaacaca	gatgtcgtcc	ccaaaatgcc	agttcaagac	4080
accttctccc	tgcccccttg	gtagtaacag	tcagggcctg	gtctgtgctc	aggtactggg	4140
tcccagtctg	ggactctgct	gctgaagttg	ccacagtaga	ggccctgggc	ttagtcctta	4200
tctccctacg	gggcttgcc	tggttttcag	tcttctctct	ctttctctct	tttttttttt	4260
tttgccacat	tctgcccctt	cctgacccca	ttgtaataac	caactccata	tccaaagggg	4320
gggtggtgct	tcagccattg	tagaagatgg	tggctttaac	ctgactgtct	aaaaattccc	4380

agctaagcct tttcctctac tctcttcctt gttctgaatc atttcttctt ctcaggccaa	4440
agtagccatg gtaaggaggc ttcattggggc agaccttgaa agatcaaaac tgcatttgca	4500
aagccctccc ctgtcccagg acaaagctga gactgacggg tgatgttgct cataggctcc	4560
agctctgcat aagaccttgg cttggagacc tccctctcag tcaacagctg aactctgagc	4620
ttgtgcccag aaattacccc aagaccacag gaaccttca agaagctccc atcacaagct	4680
tggcattgct ctctgccaca cgtgggcttc ctcaggcttg tctgccacaa gctacttctc	4740
tgagctcaga aagtggcccct tgatgaggga aaatgtccca ctgcactgcg aatttctcag	4800
ttccatttta cctcccagtc ctccttctaa accagttaat aaattcattc cacaagtatt	4860
tactgattac ctgcttgctg cagggactat tctcaggctg aagaagggtg gaggggaggg	4920
cggaacctga ggagccacct gagccagctt tatatttcaa ccatggctgg cccatctgag	4980
agcatctccc cactctcgcc aacctatcgg ggcatagccc agggatgccc ccaggcggcc	5040
caggttagat gcgtcccttt ggcttgctag tgatgacata caccttagct gcttagctgg	5100
tgctggcctg aggcagggca ggaaatcaga atagcatttg cttctctggg caaatgggaa	5160
gttcagcggg gcagcagaat cagtggcatt cccctgggtg caggccggtg ggtccactcc	5220
aactccccct gagtgtagca gcacactttc catacaccag gttctttcta caatcctggt	5280
ggaaaagcca cagaaccttc ttcctgccct tcttgagagt tccccctctt tctgggtcaa	5340
gagctggagt ggtggctcca tcctctcttg gccacttcgg tctaggaact catctttgca	5400
ggaaccagga gtcctgagca cactgaacac acctcagagg gaggatcctt gttgtggatt	5460
ttgcacctgg ctttggggca ggggtgaagt gaccaggctt agcttggtga gtttatgggc	5520
caccagggtt tggggaaatc accatcccgc ggatgctgtg acctcccttc tacggagatg	5580
caggcagtgc cagcaggggag gaggggacct gcaaagctag aatctagggc actgtttcct	5640
ccccatcctt ctctttgtag agaatagaga cgtttgctct gtctgtcttc aacctacttt	5700
tccttttctc ttttttgttt ctcatcctct ctgtgccacc tctccacca ggaggccatg	5760
tagcatagtg gaaaaagtcc ctgagggcgg ttaggagttc tgggtgacca tcctggctca	5820
gttcctaact caccatgtga catcaggcta tccccattcc cctcttggtg cctcagtttc	5880
ccgacttgca aaataagcag aaagaaccag atgctctcca ggggtctttt ctactttgct	5940
atctcatggg tcttcatttt ctcttatttt gttttctctg gatcttttcc atctgagggt	6000
acaggaagta ccaggacctg tttcagtttt tgaatcctgc aagcacattc caagactggc	6060
ctgaaactgc atgagcaaca tcaactgaaa taattttttt ttcaaaaagc accttaacaa	6120
ccaattgcga tgctgtcctg ttccttttta ctcacaccct tctctccttt ctcgtcccca	6180
tgctcccca cctcagtgtc ccgtgctgta tgcgtgtgct ctctgttctt gtataactcaa	6240
tataagtga ataaatgtgt ttgatgctga accata	6276

<211> 1209

<212> DNA

<213> NM\_173834.2| Homo sapiens hypothetical protein MGC21416 (MGC21416), mRNA

```
<400> 69
gacccgggag aaggagggcc aagatggcgg aagcggagga gtctccagga gacccgggga    60
cagcatcgcc caggcccctg tttgcaggcc tttcagatat atccatctca caagacatcc    120
ccgtagaagg agaaatcacc attcctatga gatctcgcat ccgggagttt gacagctcca    180
cattaaatga atctgttcgc aataccatca tgcgtgatct aaaagctgtt gggaaaaaat    240
tcatgcatgt tttgtacca aggaaaagta atactctttt gagagattgg gatttgtggg    300
gccctttgat ctttgtgtg acactcgcat taatgctgca aagagactct gcagatagtg    360
aaaaagatgg agggcccaa tttgcagagg tgtttgtcat tgtctggttt ggtgcagtta    420
ccatcacct caactcaaaa cttcttgagg ggaacatata ttttttcag agcctctgtg    480
tgctgggtta ctgtatactt cccttgacag tagcaatgct gatttgccgg ctggtacttt    540
tggtgatcc aggacctgta aacttcattg ttcggctttt tgtggtgatt gtgatgtttg    600
cctggtctat agttgcctcc acagctctcc ttgctgatag ccagcctcca aaccgcagag    660
ccctagctgt ttatcctggt ttctgtttt actttgtcat cagttggatg attctcacct    720
ttactcctca gtaaatcagg aatgggaaat taaaaaccag tgaattgaaa gcacatctga    780
aagatgcaat tcaccatgga gctttgtctc tggcccttat ttgtctaatt ttggagggtat    840
ttgataactg agtaggtgag gagattaaaa gggagccata tagcactgtc accccttatt    900
tgaggaactg atgtttgaaa ggctgttctt ttctctctta atgtcatttc tttaaaaata    960
catgtgcata ctacacacag tatataatgc ctcttaagg catgatggag tcaccgtggt   1020
ccatttgggt gacaaccagt gacttgggaa gcacatagat acatcttaca agttgaatag   1080
agttgataac tattttcagt tttgagaata ccagttcagg tgcagctctt aaacacattg   1140
ccttatgact attagaatat gcctctcttt tcataaataa aaatacatgg tctaaaaaaa   1200
aaaaaaaaa                                     1209
```

<210> 70

<211> 5249

<212> DNA

<213> NM\_015352.1| Homo sapiens protein O-fucosyltransferase 1 (POFUT1), transcript variant 1, mRNA

```
<400> 70
cttccctccc cgactgtgcg ccgcggctgg ctccgggttcc cgggccgaca tgggcgccgc    60
cgcgtagggca cggccgctga gcgtgtcttt cctgctgctg cttctgccgc tcccggggat   120
```

gcctgcgggc	tcctgggacc	cggccgggta	cctgctctac	tgcccctgca	tggggcgctt	180
tggaaccag	gccgatcact	tcttgggctc	tctggcattt	gcaaagctgc	taaaccgtac	240
cttggctgtc	cctccttggg	ttgagtacca	gcatcacaag	cctcctttca	ccaacctcca	300
tgtgtcctac	cagaagtact	tcaagctgga	gcccctccag	gcttaccatc	gggtcatcag	360
cttggaggat	ttcatggaga	agctggcacc	caccactggg	ccccctgaga	agcgggtggc	420
atactgcttt	gagggtggcag	cccagcgaag	cccagataag	aagacgtgcc	ccatgaagga	480
aggaaacccc	tttggcccat	tctgggatca	gtttcatgtg	agtttcaaca	agtcggagct	540
ttttacaggc	atttccttca	gtgcttccca	cagagaacaa	tgagaccaga	gattttctcc	600
aaaggaacat	ccggtgcttg	ccctgccagg	agccccagcc	cagttccccg	tcctagagga	660
acacaggcca	ctacagaagt	acatggtatg	gtcagacgaa	atggtgaaga	cgggagaggc	720
ccagattcat	gcccaccttg	tccggcccta	tgtgggcatt	catctgcgca	ttggctctga	780
ctggaagaac	gcctgtgcc	tgctgaagga	cgggactgca	ggctcgcact	tcattggcctc	840
tccgcagtgt	gtgggctaca	gccgcagcac	agcggccccc	ctcacgatga	ctatgtgcct	900
gcctgacctg	aaggagatcc	agagggctgt	gaagctctgg	gtgaggctgc	tggaatgccc	960
gtcggctctac	gttgctactg	attccgagag	ttatgtgcct	gagctccaac	agctcttcaa	1020
agggaaaggtg	aagggtggtga	gcctgaagcc	tgagggtggc	caggctcgacc	tgtacatcct	1080
cggccaagcc	gaccacttta	ttggcaactg	tgtctcctcc	ttcactgcct	ttgtgaagcg	1140
ggagcgggac	ctccagggga	ggccgtcttc	tttcttcggc	atggacaggc	cccctaagct	1200
gcgggacgag	ttctgattct	ggccggagca	ccagaccctc	tgatcctgga	gggaccagag	1260
tctgagctgg	tccttccagc	caggcctggc	agccagagggt	gctccgggat	tgcaaactcc	1320
tctttctacc	tgccaaagat	ggagaagagt	gccagggacc	cctcaaggag	ggagacgctc	1380
catatcccag	ggcataggac	ttgcagggtc	ctaggagcag	gagcatctcc	catcgcacgt	1440
gctttctgct	cttctgggaa	tttctcacac	tggaagca	gtccagcctc	cgtcttctgg	1500
tccactctgc	tctgagcagc	ctgggatgct	gaactcttca	gagagatttt	tttatagaga	1560
gattttctata	attttgatac	aagggtcatga	ctatcctaga	actctctgtg	gtttttgaaa	1620
atcattgaat	tctattaatg	taggtaccta	aagtgcctt	aactgaatgt	ggatgaggct	1680
ggggctgggtg	tggggtctttt	ggctgctttt	caagggtgtcc	cccaatgtgg	ccctcaagag	1740
ccatccccac	tgccctggcca	gagccattgt	tgtcccctac	ttcctaggcc	atttctgggg	1800
cttgggggat	gaatgctgtc	ctgtgctgta	aacactatgc	aaatggaagt	tatcggttgt	1860
gggtgctgtgc	agcgctctgt	gggcgactaa	gtgccactca	cgagcatgt	tcctggcaag	1920
gagcacatac	catcaagcca	cactatcatg	gtattgttct	cacagtcttt	tggtggttga	1980
tgccactgac	aaactgggca	ccatcagatc	tcttctgata	tcttgcccca	gtggggcctg	2040
gttggtagaa	tggtggcatt	cggttgatat	ccaaagcctg	ttctcccagc	cgtcctcctg	2100

cagctggagc	cttcaggccg	tattctcacg	agggaaacgtt	tgccaaggct	ctgacctcac	2160
agaagatgcc	cagggccag	aagccatcag	aattatcagt	ggagaagcac	cttttgactc	2220
ttcccttcca	atgtaatctc	tgccaacacc	atgaggctta	aggtgctcta	agtcatgagt	2280
gttttggctt	caaagtctgc	agttttaata	atctgtgact	cctgagagcc	catgggtttt	2340
tgaccttggt	gttctaaaat	tccttgtctg	acccctgtag	atcttttcct	tgccatgtca	2400
cctcccttgg	cctttgatcc	tggaaagggt	gcagagcctc	cactgagcca	ggcccagagc	2460
tccttgcagt	gccttcttcc	ttgtttacct	gtgggaggaa	acactttttt	tgtcaggggc	2520
agcctgggtc	agagctcaga	ggtcacactg	tatcaaagat	ctcaaacagc	aaagtcagca	2580
tttgctgtat	agagctgcca	cccaactcta	agcaggagaa	actgtacaga	aagggttttg	2640
ctatttttcc	cttttgggaa	aacaatgaag	tgttttaagt	cctgggttga	ctgagagatg	2700
gtttgcctgt	ccagacttgc	tctcaagcct	catccagaga	aggagctgca	gatgaggagg	2760
cccgtacact	ccctgccacc	actaggttgt	aagcctgtag	ctggctggct	gatttcatat	2820
tgggaattcat	ttgccatcca	cagccttaca	ctaggcacac	acttttagagt	ctggggctcc	2880
agtggggccc	gcctaatttt	ttttccccc	aagacagggc	cttgctctgt	ctcccaggct	2940
ggagtgcagt	ggcatgatca	tggcttactg	cagccttgat	ctcccaggct	caagcgatcc	3000
ttctgcctca	gcctctctgg	tagctgagac	tgcatgccca	gctccaaatc	accttgattc	3060
atatcagcag	taataatcac	ttgtgttctg	aaagaaagg	caccagaagt	tctagcaaaa	3120
ttcagttgtg	ttctgtgagc	tagcactttt	tcctctgacc	caattttctt	acctataaaa	3180
tgggtgataaa	aaccgacagg	ttgttcaaag	gccagatca	gctaaagcat	gtatataaga	3240
gcacgttgta	aacttgaaa	agacaaaggc	acaaatgtgg	ctgttgatta	atttgactgc	3300
ttctcgttgc	tcgtcacctc	catgccaggc	actgtgcttg	ctaattgctt	tatgggggca	3360
ttctcttatt	tattccccag	ccctgggaaa	taggagctgt	cattatcctt	ctctttctgc	3420
acaaggaaaa	attaatgccc	tgagaattgt	cataattttc	ccaaggctgc	ccagctgggtg	3480
gtgttaagcc	agaatttgac	ctcccagagc	cagtttccat	tagctgccat	gctctgctgc	3540
ctctaattca	cagaatgcac	tttctaccct	gtgtgccatg	gagacctcct	atggaaaaat	3600
gatcagccac	cttaccttct	actgggtacc	tgctgtgagt	ctgcctatgc	cagaaggatt	3660
aaggagggga	ggttacccaa	gaaacaaagc	ctacatgccg	cttacagccc	ccgttggtatg	3720
gttgctcagt	acaacagtct	tgcatcagc	aggtgtttgt	tcatcaccta	ctatgtgtca	3780
ggctctatgc	taggtactgg	ggatacagga	gagaatcaag	cgtaaagtct	ttgttctcaa	3840
ggaatttgca	ttctagaaa	tagaagatgt	aataaatgta	ctgtgggaca	tgtaataaag	3900
tgctataaag	aaatataaag	ggtttgggag	caaaaagagg	gagtggatct	attttagatg	3960
agcccaggta	agacctctct	gaagagctgt	catgaaggag	ggaggaggca	cattcctggc	4020
agagaaaaca	gcacgtgcaa	aggccccag	actggagtgt	gttcctgaag	agcagccagg	4080
aggccagcat	ggctggagag	gcaggcatag	gcagggaacc	gagcagcagg	tcagagcagg	4140

cgagctgaca	ttctgcagcc	tggacggcca	tggcaggaag	cttttagttg	gagagataca	4200
ggaagcctcc	tagggttctg	agcagaagag	gggcatgagc	tgattcacat	tctgaaggac	4260
ctctctagct	ggccagtgct	gaggaggttg	gagagagaaa	gggtgaaagc	agagagacca	4320
gtgcagggct	gttaacaggg	ttgcaggcga	gagactgggg	tgctgggctc	ccctagacta	4380
ggactccagt	gcccctcctc	ccaagagac	aaaggccatt	gcattgaagg	aggtgggaaa	4440
tgattagatt	ctgaacatat	gtaattat	ttcagtcctt	ttcaaagata	caaataattta	4500
catagtttta	atcatgtaat	atatacaatt	taatgtccta	gtgttttact	taatagtgtta	4560
tcatgttttc	cctgttggtg	tgtagcctgg	ataaatgctc	ttaattataa	aaaatttctgt	4620
cgaggagtgt	tccatagttt	attgttttcc	tattatgaga	atttaggcca	agtgtggtgg	4680
ctcatgcctg	taatcccagc	actttgagag	gccgaggtgg	gcagatcact	tgaggtgagg	4740
agttcaagac	cagcctggcc	aacatggtga	attatctcta	ctaaaaatac	aaaaaaataa	4800
taataatagc	caggcgtggt	ggcacatgcc	tgtattccca	gctgcttggg	aggctgaggg	4860
aggagaatgg	cttgaacctg	ggaggtggag	gttgacagtg	gccgagatgg	tgccactgca	4920
ttccagcctg	ggcaacagag	cgagactcca	tctcaaaaaa	aaggagactt	catgtgcccc	4980
caatttttca	ctattgttat	ttgaaaaaat	atTTTTtatt	gtaagagttt	ttcttttatt	5040
aaaatgttca	ttaataaagt	tgttgagcgg	gaagcaaaaa	aaaaaagttg	tttaagataa	5100
attcccagaa	gtgaatttgt	tagatcaaac	acttaaaact	ttttgttatg	gaagaattca	5160
aatataaata	aaaaattgtg	agtaataaaa	tgaactcaca	gtttcaacaa	tgaccacaca	5220
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				5249

<210> 71

<211> 722

<212> DNA

<213> NM\_175617.2| Homo sapiens metallothionein 1E (functional) (MT1E), mRNA

<400> 71						
cttgttcgtc	tcactggtgt	gagctccagc	atcccccttg	ctcgaaatgg	accccaactg	60
ctcttgcgcc	actggtggct	cctgcacgtg	cgccggctcc	tgcaagtgca	aagagtgcaa	120
atgcacctcc	tgcaagaaga	gtgagtgcgg	ggccatctcc	aggaatctgg	ggctgtggct	180
caggttggga	gggaactcaa	ggctggccct	gagtgcaccc	ttctggggaa	ctgggctttc	240
tttgccctca	ttgcccgtgt	cattccctct	ccaggctttc	tgccctaaat	tcagatgggg	300
caggacagca	tttttctcgt	gggacacaaa	ccccaaactg	accccttatg	gtttcagaac	360
agagctgtgc	cagacgaaaa	aaagcaccct	ctgggtctgg	gttctgagct	cgagccaggc	420
ttgctattag	ggcagggagg	tgcccgggtc	agtctactgc	cacctctcac	tctccccttc	480

ttccccaggc tgctgttcct gctgccccgt gggctgtgcc aagtgtgccc agggctgcgt	540
ctgcaaaggg gcatcgga ga agtgcagctg ctgtgcctga tgtgggaaca gctcttctcc	600
cagatgtaaa tagaacaacc tgcacaacct ggattttttt aaaaatacaa cactgagcca	660
tttgctgcat ttctttttat actaaatatg tgactgacaa taaaaacaat tttgacttta	720
aa	722

<210> 72

<211> 980

<212> DNA

<213> NM\_003283.3| Homo sapiens troponin T1, skeletal, slow (TNNT1), mRNA

<400> 72	
agcaaggctc agcctcaaga ttcacagcat ctcagacgca gcctaggccg caccaggatg	60
tcggacaccg aggagcagga atatgaggag gagcagccgg aagaggaggc tgcggacgag	120
gaggaggaag cccccgaaga gccggagccg gtggcagagc cagaagagga acgccccaaa	180
ccaagccgcc ccgtggtgcc tcctttgatc ccgccaaaga tcccagaagg ggagcgcgtt	240
gacttcgatg acatccaccg caagcgcgatg gagaaagacc tgctggagct gcagacactc	300
atcgatgtac atttcgagca gcggaagaag gaggaagagg agctggttgc cttgaaggag	360
cgcattgagc ggcgccggtc agagagagcc gagcaacagc gcttcagaac tgagaaggaa	420
cgcgaacgtc aggctaagct ggcggaggag aagatgagga aggaagagga agaggccaag	480
aagcgggcag aggatgatgc caagaaaaag aagggtgctgt ccaacatggg ggcccatttt	540
ggcggctacc tggtaaggc agaacagaag cgtggttaagc ggcagacggg gcgggagatg	600
aagggtgcgca tcctctccga gcgtaagaag cctctggaca ttgactacat gggggaggaa	660
cagctccggg cccggtctgc ctggctgcct ccatcacagc cctcctgccc tgccaggag	720
aaagcccagg agctgtcgga ctggatccac cagctggagt ctgagaagtt cgacctgatg	780
gcgaagctga aacagcagaa atatgagatc aacgtgctgt acaaccgcat cagccacgcc	840
cagaagttcc ggaagggggc aggggaagggc cgcgttgagg gccgctggaa gtgaggatgc	900
cgccccggac agtggcacct gggaagcctg ggagtgtttg tcccatcggt agcttgaaat	960
aaacgctccc ctcagacacc	980

<210> 73

<211> 2213

<212> DNA

<213> NM\_004067.1| Homo sapiens chimerin (chimaerin) 2 (CHN2), mRNA

<400> 73  
gggcgtgcaa aggcgcggag cgggacggaa accacaaata aatagcggcg gcggcagcgc 60  
gtcatctggt ggagcaggaa gtgcaggcag agtccggagg ctggtgcttt ctgcgcgtcc 120  
ccaggacttt gccatgggct gggggccgcg gaggctgcga gcggccgggc gagggcagcg 180  
gcggcggcgt cccacccggg gctgagcgag cagcgacgcg aggggcgcgc ggagatggca 240  
gcgtccagca actccagcct gtccggctcg tcggtgtcct ccgatgctga agaataccag 300  
cctcctatat ggaaatcata cttatatcag ttacagcaag aggcacctcg tccaagaga 360  
atcatttgtc ctccggagggt ggaaaacaga ccaaaatatt atggaagaga gtttcatggg 420  
atcatctctc gggagcaggc ggatgagctt cttggaggcg tggaggggtgc ctacatcctt 480  
agagaaagcc agcggcaacc aggatgctac acgctggctc tcagggtttgg aaaccagacc 540  
ttaaactaca ggctcttcca cgacgggaaa cactttgtgg gtgagaagag gtttgagtcg 600  
attcatgata tgggtgacaga tggcttgata acactgtaca tagaaacaaa agctgccgag 660  
tacatttcaa aaatgacaac taaccccatc tatgaacaca ttggatatgc caccctactc 720  
agagaaaaag tatccagaag gctgagcagg tctaaaaatg aaccaagaaa aacaaacgtc 780  
acacatgaag aacacacagc ggtggaaaag atctcctccc tggttcgaag ggctgccctc 840  
acacacaacg acaaccactt caattatgag aagacacaca actttaaggt ccacacgttc 900  
cgaggcccac actggtgtga atattgtgcc aatttcatgt gggggctcat cgcccaaggg 960  
gtccggtgct cagactgtgg attgaacgta cacaacagct gttccaagca cgttcccaat 1020  
gactgccaac ctgatctcaa gaggatcaag aaagtgtact gttgtgacct cacaacactt 1080  
gtgaaggctc acaacactca gagacccatg gtggtagaca tatgcattcg ggaaattgaa 1140  
gcaagaggat taaaatcggg aggcctttac agagtctctg ggttcactga acacattgaa 1200  
gatgtcaaaa tggcatttga cagagatggt gaaaaggccg atatatctgc caatgtctat 1260  
ccagacataa acatcatcac tggagccctt aaactgtatt tcagagactt acccatccct 1320  
gtcatcacat atgataccta ttccaaattt atagatgcag caaaaatctc caatgcagat 1380  
gagaggctgg aagccgtcca tgaagtgtcg atgctgtgc ctctgcca ctatgaaacc 1440  
ctccggtacc taatgatcca cctcaaaaag gttactatga atgaaaaaga caatttcatg 1500  
aatgcagaaa atctgggggat cgtgtttggg ccactctga tgaggcccc tgaggacagc 1560  
accctgacca ccctgcatga tatgcggtac caaaagctga ttgtgcagat ttaatagaa 1620  
aacgaagacg ttttattcta atccatcagg gaaatgagct gaatggcca gcaccatcaa 1680  
gttgacacag ctaaggataa aacatttctt accacttgat ttgttttcca agcaagtgtc 1740  
agaatttgct ggactgcaga ggatcgctga gtgggggtact gtgtctcata gacatgcgcc 1800  
acctccacgt gagaacaagg gtgaagggtga ggggaagcccc tcagggttggg tcttttgctg 1860  
tgcctcctat gtatgtctgg tttgctggaa gagtgattaa tacatcttta atttattaaa 1920  
aaacaatgta gacctttaa cttcagtcctt attgggaata aaagggaact taattcatac 1980



aggtacttga tacagttata cttttccac ttacaaaaag aagacaattc tgttaaata	2040
aacgtgtatc gtaaaatgta attttattta cccacgagaa tgttggtatt ttagcaatag	2100
aactcaatgc agatgcattg gttattaccc tgtgtacctt gtccctcatt ttgctgtgac	2160
accctgaaaa agctgaccac aaatgcagta ttatcattga catacctctg tcc	2213

<210> 74

<211> 2201

<212> DNA

<213> NM\_005520.1| Homo sapiens heterogeneous nuclear ribonucleoprotein H1 (H) (HNRPH1), mRNA

<400> 74	
tttttttttt cgtcttagcc acgcagaagt cgcgtgtcta gtttgtttcg acgccggacc	60
gcgtaagaga cgatgatggt gggcacggaa ggtggagagg gattcgtggt gaagggtccgg	120
ggcttgccct ggtcttgctc ggccgatgaa gtgcagagggt ttttttctga ctgcaaaatt	180
caaaatgggg ctcaagggtat tcgtttcatc tacaccagag aaggcagacc aagtggcgag	240
gcttttggtt aacttgaatc agaagatgaa gtcaaattgg ccctgaaaaa agacagagaa	300
actatgggac acagatatgt tgaagtattc aagtcaaaca acgttgaaat ggattgggtg	360
ttgaagcata ctggtccaaa tagtcctgac acggccaatg atggctttgt acggcttaga	420
ggacttcctt ttggatgtag caaggaagaa attgttcagt tcttctcagg gttggaaatc	480
gtgccaaatg ggataacatt gccggtggac ttccagggga ggagtacggg ggaggccttc	540
gtgcagtttg cttcacagga aatagctgaa aaggctctaa agaaacacaa ggaaagaata	600
gggcacagggt atattgaaat ctttaagagc agtagagctg aagttagaac tcattatgat	660
ccaccacgaa agcttatggc catgcagcgg ccaggtcctt atgacagacc tggggctggt	720
agaggggtata acagcattgg cagaggagct ggctttgaga ggatgaggcg tggtgcttat	780
ggtggagggt atggaggcta tgatgattac aatggctata atgatggcta tggatttggg	840
tcagatagat ttggaagaga cctcaattac tgtttttcag gaatgtctga tcacagatac	900
ggggatggtg gctctacttt ccagagcaca acaggacact gtgtacacat gcggggatta	960
ccttacagag ctactgagaa tgacatttat aatttttttt caccgctcaa ccctgtgaga	1020
gtacacattg aaattggtcc tgatggcaga gtaactggtg aagcagatgt cgagttcgca	1080
actcatgaag atgctgtggc agctatgtca aaagacaaag caaatatgca acacagatat	1140
gtagaactct tcttgaattc tacagcagga gcaagcgggt gtgcttacga acacagatat	1200
gtagaactct tcttgaattc tacagcagga gcaagcgggt gtgcttatgg tagccaaatg	1260
atgggaggca tgggcttgct aaaccagtcc agctacgggg gccagccag ccagcagctg	1320
agtggggggt acggaggcgg ctacggtggc cagagcagca tgagtggata cgaccaagtt	1380

ttacaggaaa	actccagtga	ttttcaatca	aacattgcat	aggtaaccaa	ggagcagtga	1440
acagcagcta	ctacagtagt	ggaagccgtg	catctatggg	cgtgaacgga	atgggaggggt	1500
tgtctagcat	gtccagtatg	agtggtggtg	ggggaatgta	attgatcgat	cctgatcact	1560
gactcttggt	caaccttttt	tttttttttt	ttttctttaa	gaaaacttca	gtttaacagt	1620
ttctgcaata	caagcttggtg	atztatgctt	actctaagtg	gaaatcagga	ttgttatgaa	1680
gacttaaggc	ccagtatttt	tgaatacaat	actcatctag	gatgtaacag	tgaagctgag	1740
taaactataa	ctgttaaact	taagttccag	cttttctcaa	gttagttata	ggatgtactt	1800
aagcagtaag	cgtattttagg	taaaagcagt	tgaattatgt	taaatggtgc	cctttgccac	1860
gttaaattga	acactgtttt	ggatgcatgt	tgaaagacat	gcttttattt	tttttgtaaa	1920
acaatatagg	agctgtgtct	actattaaaa	gtgaaacatt	ttggcatggt	tgtaatttct	1980
agtttcattt	aataacctgt	aaggcacgta	agttaagct	tttttttttt	ttaagttaat	2040
gggaaaaaatt	tgagacgcaa	taccaatact	taggattttg	gtcttggtgt	ttgtatgaaa	2100
ttctgaggcc	ttgattttaa	tctttcattg	tattgtgatt	tccttttagg	tatattgcgc	2160
taagtgaaac	ttgtcaaata	aatcctcctt	ttaaaaactg	c		2201

<210> 75

<211> 1895

<212> DNA

<213> NM\_004046.4| Homo sapiens ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle (ATP5A1), nuclear gene encoding mitochondrial protein, transcript variant 2, mRNA

<400> 75						
gtcttgacct	tctttgcggc	tcggccattt	tgtcccagtc	agtccggagg	ctgcggctgc	60
agaagtaccg	cctgcggagt	aactgcaaag	atgctgtccg	tgcgcgttgc	tgcggccgtg	120
gtccgcgccc	ttcctcggcg	ggccggactg	gtctccagaa	atgctttggg	ttcatctttc	180
attgctgcaa	ggaacttcca	tgcttctaac	actcatcttc	aaaagactgg	gactgctgag	240
atgtcctcta	ttcttgaaga	gcgtattctt	ggagctgata	cctctgttga	tcttgaagaa	300
actgggcgtg	tcttaagtat	tggtgatggt	attgcccgcg	tacatgggct	gaggaatggt	360
caagcagaag	aaatggtaga	gttttcttca	ggcttaaagg	gtatgtcctt	gaacttggaa	420
cctgacaatg	ttggtgttgt	cgtgttttga	aatgataaac	taattaagga	aggagatata	480
gtgaagagga	caggagccat	tgtggacggt	ccagttgggtg	aggagctggt	gggtcgtgta	540
gttgatgccc	ttggtaatgc	tattgatgga	aagggtccaa	ttggttccaa	gacgcgtagg	600
cgagttgggtc	tgaaagcccc	cggtatcatt	cctcgaattt	cagtgcggga	accaatgcag	660
actggcatta	aggctgtgga	tagcttgggtg	ccaattgggtc	gtggtcagcg	tgaactgatt	720
attggtgacc	gacagactgg	gaaaacctca	attgctattg	acacaatcat	taaccagaaa	780

cgtttcaatg atggatctga tgaaaagaag aagctgtact gtatttatgt tgctattggt	840
caaaagagat ccactgttgc ccagttggtg aagagactta cagatgcaga tgccatgaag	900
tacaccattg tgggtgctggc tacggcctcg gatgctgccc cacttcagta cctggctcct	960
tactctggct gttccatggg agagtatttt agagacaatg gcaaacaatgc tttgatcatc	1020
tatgacgact tatccaaaca ggctgttgct taccgtcaga tgtctctgtt gctccgccga	1080
ccccctggtc gtgaggccta tcctggtgat gtgttctacc tacactcccg gttgctggag	1140
agagcagcca aaatgaacga tgcttttggt ggtggctcct tgactgcttt gccagtcata	1200
gaaacacagg ctggtgatgt gtctgcttac attccaacaa atgtcatttc catcactgac	1260
ggacagatct tcttggaac agaattgttc tacaaaggta tccgccctgc aattaacgtt	1320
ggtctgtctg tatctcgtgt cggatccgct gcccaaacca gggctatgaa gcaggtagca	1380
ggtaccatga agctggaatt ggctcagtat cgtgaggttg ctgcttttgc ccagttcggg	1440
tctgacctcg atgctgccac tcaacaactt ttgagtcgtg gcgtgcgtct aactgagttg	1500
ctgaagcaag gacagtattc tcccatggct attgaagaac aagtggctgt tatctatgcg	1560
ggtgtaaggg gatattctga taaactggag ccagcaaga ttacaaagtt tgagaatgct	1620
ttcttgtctc atgtcgtcag ccagcaccaa gccttggtgg gcactatcag ggctgatgga	1680
aagatctcag aacaatcaga tgcaaagctg aaagagattg taacaaattt cttggctgga	1740
tttgaagctt aaactcctgt ggattcacat caaataccag ttcagttttg tcattgttct	1800
agtaaattag ttccatttgt aaaaggggta ctctcatact ccttatgtac agaaatcaca	1860
tgaaaaataa aggttcata atgcatagtt aaaaa	1895

<210> 76

<211> 1290

<212> DNA

<213> NM\_001970.3| Homo sapiens eukaryotic translation initiation factor 5A (EIF5A), mRNA

<400> 76	
gcggcggcgg cggtagaggc ggcggcggcg gcggcagcgg gctcggaggc agcggttggg	60
ctcgcggcga gcggacgggg tcgagtcagt gcgttcgcgc gagttggaat cgaagcctct	120
taaaatggca gatgacttgg acttcgagac aggagatgca ggggcctcag ccaccttccc	180
aatgcagtgc tcagcattac gtaagaatgg ctttgtggtg ctcaaaggcc ggccatgtaa	240
gatcgtcag atgtctactt cgaagactgg caagcacggc cagccaagg tccatctggt	300
tggtattgac atctttactg ggaagaaata tgaagatatc tgcccgtcaa ctcataatat	360
ggatgtcccc aacatcaaaa ggaatgactt ccagctgatt ggcattcagg atgggtacct	420
atcactgctc caggacagcg gggaggtacg agaggacctt cgtctccctg agggagacct	480

tggaaggag attgagcaga agtacgactg tggagaagag atcctgatca cggtgctgtc	540
tgccatgaca gaggaggcag ctgttgcaat caaggccatg gcaaaataac tggctcccag	600
gatggcgggtg gtggcagcag tgatcctctg aacctgcaga ggccccctcc ccgagcctgg	660
cctggctctg gcccggctct aagctggact cctcctacac aattttattg acgtttttatt	720
ttggttttcc ccaccccctc aatctgtcgg ggagccccctg cccttcacct agctcccttg	780
gccaggagcg agcgaagctg tggccttggg gaagctgccc tcctcttctc ccctcacact	840
acagccctgg tgggggagaa ggggggtggg gctgcttgtg gtttagtctt tttttttttt	900
tttttttttt ttttaaattc aatctggaat cagaaagcgg tggattctgg caaatggctc	960
ttgtgcctc ccactcctc cctggctctg tcccctgttg cccatagccc tttaccctga	1020
gcaccacccc aacagactgg ggaccagccc cctcgctgc ctgtgtctct ccccaaacc	1080
cttagatgg ggagggaaga ggaggagagg ggaggggacc tgccccctcc tcaggcatct	1140
gggagggccc tgcccccatg ggctttaccc ttccctgcgg gctctctccc cgacacattt	1200
gttaaaatca aacctgaata aaactacaag tttaatatga aaaaaaaaaa aaaaaaaaaa	1260
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1290

<210> 77

<211> 2512

<212> DNA

<213> NM\_005041.3| Homo sapiens perforin 1 (pore forming protein) (PRF1), mRNA

<400> 77	
ggctgggtgca aggagccaca gtgggctgcc tggggggctg atgccaccat tccaggagcc	60
tcggtgaaga gaggatatcc atctgtgtag ccgcttctct atacgggatt ccagctccat	120
ggcagcccgt ctgctcctcc tgggcctcct tctcctgctg ctgcccctgc ccgtccctgc	180
cccgtgccac acagccgcac gctcagagtg caagcgcagc cacaagtctg tgcctgggtgc	240
atggctggcc ggggaggggtg tggacgtgac cagcctccgc cgctcgggct ccttcccagt	300
ggacacacaa aggttcctgc ggcccgcagg cacctgcacc ctctgtgaaa atgccctaca	360
ggagggcacc ctccagcgcc tgcctctggc gtcaccaac tggcgggccc agggctctgg	420
ctgccagcgc catgtaacca gggccaaagt cagctccact gaagctgtgg cccgggatgc	480
ggctcgtagc atccgcaacg actggaaggt cgggctggac gtgactccta agcccaccag	540
caatgtgcat gtgtctgtgg ccggctcaca ctacaggca gccaaacttg cagcccagaa	600
gacccaccag gaccagtaca gcttcagcac tgacacgggt gagtgccgct tctacagttt	660
ccatgtggta cacactcccc cgctgcaccc tgacttcaag agggccctcg gggacctgcc	720
ccaccacttc aacgcctcca ccagcccgcc ctacctcagg cttatctcca actacggcac	780

ccacttcatc cgggctgttg agctgggttg ccgcataatc gccctcactg ccctgcgcac	840
ctgcgagctg gccctggaag ggctcacgga caacgaggtg gaggactgcc tgactgtcga	900
ggcccagggtc aacataggca tccacggcag catctctgcc gaagccaagg cctgtgagga	960
gaagaagaag aagcacaaga tgacggcctc cttccaccaa acctaccggg agcgccactc	1020
ggaagtgggtt ggcggccatc acacctccat taacgacctg ctgttcggga tccaggccgg	1080
gcccagagcag tactcagcct gggtaaaactc gctgcccggc agccctggcc tgggtggacta	1140
caccctggaa cccctgcacg tgctgttgga cagccaggac ccgcgccggg aggcactgag	1200
gagggccctg agtcagtacc tgacggacag ggctcgctgg agggactgca gccggccgtg	1260
cccaccaggg cggcagaaga gcccccgaga cccatgccag tgtgtgtgcc atggctcagc	1320
ggtcaccacc caggactgct gccctcggca gaggggcctg gccagctgg aggtgacctt	1380
catccaagca tggggcctgt ggggggactg gttcactgcc acggatgcct atgtgaagct	1440
cttcttttgtt ggccaggagc tgaggacgag caccgtgtgg gacaataaca accccatctg	1500
gtcagtgcgg ctggattttt gggatgtgct cctggccaca ggggggcccc tgaggttgca	1560
ggtctgggat caggactctg gcagggacga tgacctcctt ggcacctgtg atcaggctcc	1620
caagtctggt tcccatgagg tgagatgcaa cctgaatcat ggccacctaa aattccgcta	1680
tcatgccagg tgcttgcccc acctgggagg aggcacctgc ctggactatg tcccccaaat	1740
gcttctgggg gagcctccag gaaaccggag tggggccgtg tgggtgagaac agtgagcttg	1800
gaaaggacca gtatgcttg actgaagggg ttctcacagt gggagccagg gctgtcttcg	1860
tattcccatt agaccaagct tgtccaacct gagggccgca tgcggcccag gatggctttg	1920
aatgcccggc aacgcaaatt cgcaaacttt cttaaaacat tatgagtttc tttttgctat	1980
tttttttttt ttttttagctc atcggctatc gttagtgtca gtggatttta catgtggccc	2040
aacacaattc ttcttccaac gtggcccaga gaagccaaaa gattggatac gcatcagaca	2100
gatggaaaag ggagattcag actgtttttc agggaggttg ctgggtttac acgctaattc	2160
cgattcaccc tgtccaaact gcctaagccc tccgccattc tcaagccctg cagtcacagc	2220
tacacagatc acagcttcag ccaggagctg ggcagaaggc caagaggctg ttcccaccag	2280
gctgtcagg gctgggtcttt taggacctt cccttgagcc ctctatggtg tggcaaagcc	2340
ttcattgcct taactggagc cccatcagct ccagctgtc tgtcttcttt gccacaatg	2400
ctttgcccct gagacaaatg gaggcctgtc ctgacctgtc tcaccatgta catagcttga	2460
taaagggcca ataaatatga tgttatggtg aaaaaaaaaa aaaaaaaaaa aa	2512

<210> 78  
 <211> 4623  
 <212> DNA

<213> NM\_014965.2| Homo sapiens OGT(O-Glc-NAC transferase)-interacting protein  
106 KDa (OIP106), mRNA

```

<400> 78
gatgctgggc caggagcttt gtgtacaccc ctccacttca gctgagccag ggcattgtctg    60
cgccccaggc cagggcgagc tgtgtgccct gggggcccag gcctgcatgg ctccctctggg    120
tagggggctc ggggcacccc caaggatggg cccttagggg gatgttttgg ctttgggggtg    180
acttcagcaa tgtccctgcg agacaagggc ggggaagaag aatgttttga atacgactgc    240
caggatgaag agaggaagcc aaccacagg cagcatgaca ccaggacct cttggaagag    300
gttttatgtg ctgaaagagt tggccagatg actaagacat ataatgacat agatgctgtc    360
actcggcttc ttgaggagaa agagcgggat ttagaattgg ccgctcgcag cggccagtcg    420
ttgttgaaga agaacaagac cctaaccgag aggaacgagc tgctggagga gcaggaggaa    480
cacatcaggg aggaggtgtc tcagctccgg catgagctgt ccatgaagga tgagctgctt    540
cagttctaca ccagcgtgc ggaggagagt gagcccagat ccgtttgctc aaccccgttg    600
aagaggaatg agtcgtcctc ctcatgccag aattactttc atttggattc tcttcaaaag    660
aagctgaaag accttgaaga ggagaatgtt gtacttcgat ccgaggccag ccagctgaag    720
acagagacca tcacctatga ggagaaggag cagcagctgg tcaatgactg cgtgaaggag    780
ctgagggatg ccaatgtcca gattgctagt atctcagagg aactggccaa gaagacggaa    840
gatgctgccc gccagcaaga ggagatcaca cacctgctat cgcaaatagt tgatttgag    900
aaaaaggcaa aagcttgccg agtggaaaat gaagaacttg tccagcatct gggggctgct    960
aaggatgccc agcggcagct cacagccgag ctgcgtgagc tggaggacaa gtacgcagag   1020
tgcatggaga tgctgcatga ggcgaggag gagctgaaga acctccggaa caaaaccatg   1080
cccaatacca cgtctcggcg ctaccactca ctgggcctgt ttcccatgga ttccttggca   1140
gcagagattg agggaacgat gcgcaaggag ctgcagttgg aagaggccga gtctccagac   1200
atcactcacc agaagcgtgt ctttgagaca gtaagaaaca tcaaccaggt tgtcaagcag   1260
agatctctga ccccttctcc catgaacatc cccggctcca accagtcctc ggccatgaac   1320
tccctcctgt ccagctgcgt cagcaccccc cggctccagct tctacggcag cgacataggg   1380
aacgtcgtcc tcgacaacaa gaccaacagc atcattcttg aaacagaggc agccgacctg   1440
ggaaacgatg agcggagtaa gaagccgggg acgccgggca cccaggctc ccacgacctg   1500
gagacggcgc tgaggcggct gtccctgcgc cgggagaact acctctcgga gaggagggtc   1560
tttgaggagg agcaagagag gaagctccag gagctggcgg agaagggcga gctgcgcagc   1620
ggctccctca caccactga gagcatcatg tccctgggca cgcactcccg cttctccgag   1680
ttcaccggct tctctggcat gtccttcagc agccgctcct acctgcctga gaagctccag   1740
atcgtgaagc cgctggaagg ttccgccaca cttcaccact ggcagcagtt ggcccaacct   1800
caccttgggg gcatactgga cccccggccc ggtgtggtca ccaagggctt ccggacgctg   1860

```

gatgttgacc tggacgaagt gtactgcctt aacgactttg aagaagatga cacaggtgac	1920
cacatttctc tcccacgcct agctacctcc actccagttc agcaccacaga gacctcaggt	1980
gagaggtccc aagcacgtgt gactgtctca ggcagcagaa gttacccgag ccggcctcag	2040
gcttccccag aggagatgca ggagccgcca gcggccacgg aggaggagga ggaggaggag	2100
gaggaggagg ggtctggtga gggcaccacg ataagtctg taaacttggc acctttcccg	2160
gaggcagagt tttgggccat tctcacctct gttccaggca ccatccgtag tggttctctg	2220
tctgtagctt ccgctcgtct gtgtgggtga tgattaaagc attctcattg cacagttctg	2280
tttttaaata cagagtctga tgcctcctat ttgtaacaat ggggtgtagct cccctgcccc	2340
tcttgagggt gcatggcccc tcagggatct ttaaagtggg agcaggaaag gctgctaaaa	2400
aaaaaaaaaa aaaaagtggg cttttgggtc cctgaaaaca tcagtgcctt tcttcctggt	2460
ctgggtgtct ccctgagtct aaggggaaga ttctcaagtc ccctggtgat ttccaagtgg	2520
agctgagcag ttttagggaa attgagtgtt gggtcattca gaaggtaaat gagatcatct	2580
gttacctgta cgctgtatta aaatagaacc aggaaaggct caggatttca gacatttcgt	2640
cagccttttc actttcccag ctccaatgga ggtatatatg tcattttctt ttcagcttac	2700
acatgtgttc aaagtggatt tttaaaaagt gttttagcaa tactccttaa ccaaataaac	2760
cttcggagaa catcactaag cttttccaga gagaagacct tagatgaagt tggaaaagag	2820
ctctggctga ctccaccac tgtccccaag cattaaagggt gtggccatga gtttacagaa	2880
ttaccacat tccagttgcc actgggatga aaagctttgt gctcagagct ctgggaatca	2940
tgggatcaca tggttactgt ttacccaac agcatgtctt gtctagactg gacctccag	3000
ccccttgtgt tgggaaggca aagcttttgt ggagtcaggg gggagactga gaaaaaatga	3060
attaaccctg tgttgtcatc ctcatgacat tcctgaggat tcaccagggtt agaagtgagg	3120
acgtttatct ttgtgattca taattttcat ttgtgaaggc cacaacactc cccttgaaat	3180
acagggcagg aaagagctga ggttcttggg ggtgtctccc attcccctgg ctaatttcag	3240
acagctgtgg tcaagggatt ttacttgga tcaacttttc cttttttcct tgacattaat	3300
tttagagaag tcatcaagtc atgtgatttg tttagcacat aggcctatct atggtttgat	3360
tttttttagg cagttatatt actagattaa gcttgtgagg gaatgaaaat gttttttatt	3420
tgttatctac acactcgaaa aagagaaacc agctgcggta ctgtcccatt tttgtcatca	3480
gcaccagtgt ccgtcaggaa ggcagggtgt ggtgcagaaa catgatgcct ggctgatttt	3540
cgtggctaaa ggggtaggcc ttatgttgat tgggatgtc ccctacagcc ttacaggtag	3600
aatagaagggt gagttctgga agtgcaaaga agcaccatta agtgcatctt ctagaagttg	3660
tacagaagga ctaaagctta tcaagtcaac aaagaactta ctttgaggga tagagagaga	3720
gaagggacca attatgagga ctgacatcct ggccactctc cttaaaataa aactgacat	3780
ttttcttgct tgttctctct gtactcaaac ctgtggcaaa ttcatcctag caacgttatt	3840

tgacgagggg catgaacatt tatagttgaa actgtagaaa gggtcagggtt ggaggtgtgt	3900
aataaaaaag aattacctag gttgccaaag gtaatttagg aaggggtctga tcatttagat	3960
gagagtcttt tggggcttat tttctgggta aggtcatct ttaaaaactg gcttcagagg	4020
ggagagggga gaacaatgaa ttggctctat tttctctatt gggaattaca ggaccatttt	4080
gattcttaga atgtaaaaag catatcgcta agtaaatcat cctggagggtc ccaagtagct	4140
ctatgcctgc aatcatggag acacaggcag acagataagc ttcattgggga aggcattggg	4200
catcctctgt cttgggattt gtatccatgg tggctctggc cctgcctttt aatccgtcct	4260
ctacgcttgg gcttttctgt taccaaacag cactatccca ggaactattg tctgcctggg	4320
aacactcagt agggagacac tttggagaca ggaggtgatg aaccttttta tgtgcagctg	4380
gtatgataga aggaaattgg gaaaacttgt atgctaggca cttttgtcca gagcctgctg	4440
tcccatggag aaaaagtttt aagcactgaa aaaatttgat taatgtattt aaatgtatta	4500
tttgaagcat cattcacttg ttgattttta caatcccatg tcttaaaaag gatgaatcca	4560
tgttattgta ttgtaataa tttagattat taaaatggat tgtttaaaaa aaaaaaaaaa	4620
aaa	4623

<210> 79

<211> 2657

<212> DNA

<213> NM\_017895.6| Homo sapiens DEAD (Asp-Glu-Ala-Asp) box polypeptide 27 (DDX27), mRNA

<400> 79	
aagtgacgca tggctacttgc gcaaagacga cgaggaggct gcgaaaagtt aagggccgga	60
ccgcaggctg tgctcgcttc cggaagtggc ttctgcgaca acatgcttgc ggacctcggc	120
ttaatcgga ccataggcga ggatgacgag gtgccggtgg agcccagatc tgactccggg	180
gacgaggaag agggagggcc cattgtgctg ggcagacgac aaaaagcttt ggggaagaac	240
cgcagtgtg atttcaaccc tgatttcgtt ttcactgaga agggagggac gtacgatggc	300
agctggggcc tggctgatgt catgagccaa ctcaagaaga agagggcagc cactacatta	360
gatgagaaga ttgagaaagt tcgaaagaaa aggaaaacag aggataaaga agccaagtct	420
gggaagtgtg aaaaggagaa agaagcaaag gaaggctctg aaccaaagga gcaggaagac	480
cttcaagaga atgatgagga aggtcagaa gatgaagcct cggagactga ctactcatca	540
gctgatgaga acatcctcac caaagcagat aactcaaag taaaggatcg gaagaagaag	600
aagaagaaag gacaggaagc agggaggattt tttgaagatg catctcagta cgatgaaaac	660
ctctcgttcc aggacatgaa cttttccgc cctcttctga aggccattac agccatgggc	720
ttcaagcagc ccaccccgat ccagaaggcg tgcatacctg tgggtctatt ggggaaggac	780



atctgtgcct gtgcagccac tgggacaggt aaaactgccg cctttgccct gcctgttttg	840
gagcgtctga ttataaacc ccgccaggct ccagtcaccc gcgtgctggt gctagtgtccc	900
acccgagagc tgggcatcca ggtgcactct gtcaccagac agctggccca gttctgcaac	960
atcaccacct gcctggctgt gggcggcttg gatgtgaagt ctcaggaagc agctcttcgg	1020
gcagcgctg acatcctcat cgccaccca ggccggctca tcgatcacct ccacaactgc	1080
ccttccttcc acctgagcag catcgagggtg ctcatccttg acgaggctga caggatgctg	1140
gatgagtact ttgaggagca gatgaaggag atcatccgaa tgtgttccca ccaccgccag	1200
accatgctct tctcgccac catgacagac gaggtgaaag atctggcttc tgtctccttg	1260
aagaatcctg tccgatatt tgtgaacagc aacacagatg tggctccctt cctgcggcag	1320
gagttcatcc ggatccggcc taatcgtgaa ggagaccggg aagccatcgt ggcagctttg	1380
ttgacgagga ctttactga ccatgtgatg ctgttcacgc aaaccaagaa gcaggccac	1440
cgcattgcaca tctcctggg gctcatgggg ctgcagggtg gtgagctcca tggcaacttg	1500
tcacagacgc agcggctgga ggccctccgg cgttttaagg atgaacagat tgacatcctc	1560
gtggccactg atgtggcagc ccgtggactt gacattgagg gggtcaaaac ggtaatcaac	1620
ttcacaatgc ctaataccat caaacattat gtccaccggg tggggcgaa agcacgtgct	1680
ggcagggtg ggcgtcagt ctctctggtg ggagaagatg agcgggaagat gctgaaggag	1740
attgtaaaag ctgccaaagg cctgtgaag gccaggatac tcccccaaga tgtcatcctc	1800
aaattccggg acaagattga gaaaatggag aaagatgtgt atgcagttct gcagctagag	1860
gcggaggaaa aagagatgca gcagtcagaa gccagatca atacagcaaa gcggctcctg	1920
gagaagggga aggaggcagt ggtccaagag cccgagagga gctgggtcca gaccaagaa	1980
gagaggaaga aggagaaaat tgccaaagct ctgcaggaat ttgacttggc cttagagga	2040
aagaagaaaa ggaagaagtt tatgaaggat gccaaaaaaa agggggagat gacagcagag	2100
gaaagggtctc agtttgaaat cctcaaggcg cagatgtttg ctgaacggct agcgaagagg	2160
aatcgagag ccaagcgggc ccgagcaatg cccgaggagg agccagtgag aggtcctgcc	2220
aagaagcaaa agcaggggaa gaaatctgta ttgatgaag aactcaccaa cacaagcaag	2280
aaggccctga aacagtatcg agctagccct tcctttgaag aaaggaaaca gttgggcttg	2340
ccccaccaga gacgaggagg aaactttaaa tctaaatcca gatacaagag gaggaagtag	2400
ctgtcgtggc ctgaagaaat tcatgggggc agcccttaaa tcccttcctt gtgggaagtc	2460
atcctggctg gtctgtcttt tctccatttg tttaaaaaaa aaacaaaaac aaaaaacaac	2520
actttggtgt ggtgggtatg tacgtagcta ttttctaag catgtctgtc aatctccctt	2580
cttgctgatt agctttcata tgactatatt aaatggaagt atttttggga aaagagaaac	2640
caaaaaaaaa aaaaaaa	2657

<211> 3246

<212> DNA

<213> NM\_018206.3| Homo sapiens vacuolar protein sorting 35 (yeast) (VPS35), mRNA

<400> 80  
ctacgcgcgg ggcgggtgct gcttgctgca ggctctgggg agtcgccatg cctacaacac 60  
agcagtcctc tcaggatgag caggaaaagc tcttgatga agccatacag gctgtgaagg 120  
tccagtcatt ccaaataag agatgcctgg acaaaaacaa gcttatggat gctctaaaac 180  
atgcttctaa tatgcttggt gaactccgga cttctatgtt atcaccaaag agttactatg 240  
aactttatat ggccatttct gatgaactgc actacttga ggtctacctg acagatgagt 300  
ttgctaaagg aaggaaagtg gcagatctct acgaacttgt acagtatgct ggaaacatta 360  
tcccaaggct ttaccttttg atcacagttg gaggttgata tgtcaagtca tttcctcagt 420  
ccaggaagga tattttgaaa gatttggtag aaatgtgccg tgggtgtgcaa catcccttga 480  
ggggtctgtt tcttcgaaat taccttcttc agtgtaccag aaatatctta cctgatgaag 540  
gagagccaac agatgaagaa acaactggtg acatcagtga ttccatggat tttgtactgc 600  
tcaactttgc agaaatgaac aagctctggg tgcgaatgca gcatcagggg catagccgag 660  
atagagaaaa aagagaacga gaaagacaag aactgagaat tttagtggga acaaatattg 720  
tgcgccctcag tcagttggaa ggtgtaaatg tggaaagtta caaacagatt gttttgactg 780  
gcatattgga gcaagttgta aactgtaggg atgctttggc tcaagaatat ctcattggagt 840  
gtattattca ggttttccct gatgaatttc acctccagac tttgaatcct tttcttcggg 900  
cctgtgctga gttacaccag aatgtaaatg tgaagaacat aatcattgct ttaattgata 960  
gattagcttt atttgctcac cgtgaagatg gacctggaat cccagcggat attaaacttt 1020  
ttgatataat ttcacagcag gtggctacag tgatacagtc tagacaagac atgccttcag 1080  
aggatgttgt atctttacaa gtctctctga ttaatcttgc catgaaatgt taccctgatc 1140  
gtgtggacta tgttgataaa gttctagaaa caacagtggg gatattcaat aagctcaacc 1200  
ttgaacatat tgctaccagt agtgcagttt caaaggaact caccagactt ttgaaaatac 1260  
cagttgacac ttacaacaat attttaacag tcttgaaatt aaaacatttt caccactctt 1320  
ttgagtactt tgactacgag tccagaaaga gcatgagttg ttatgtgctt agtaatgttc 1380  
tggattataa cacagaaatt gtctctcaag accaggtgga ttccataatg aatttggtat 1440  
ccacgttgat tcaagatcag ccagatcaac ctgtagaaga ccctgatcca gaagattttg 1500  
ctgatgagca gagccttgtg ggccgcttca ttcattctgct gcgctctgag gacctgacc 1560  
agcagtactt gattttgaac acagcacgaa aacatttttg agctgggtgga aatcagcggg 1620  
ttcgcttcac actgccacct ttggtatttg cagcttacca gctggctttt cgatataaag 1680  
agaattctaa agtggatgac aaatgggaaa agaaatgccg gaagattttt tcatttgccc 1740

accagactat cagtgccttg atcaaagcag agctggcaga attgccctta agactttttc	1800
ttcaaggagc actagctgct ggggaaattg gttttgaaaa tcatgagaca gtcgcatatg	1860
aattcatgtc ccaggcattt tctctgtatg aagatgaaat cagcgattcc aaagcacagc	1920
tagctgccat caccttgatc attggcactt ttgaaaggat gaagtgcctc agtgaagaga	1980
atcatgaacc tctgaggact cagtgtgccc ttgctgcac caaacttcta aagaaacctg	2040
atcagggccg agctgtgagc acctgtgcac atctcttctg gtctggcaga aacacggaca	2100
aaaatgggga ggagcttcac ggaggcaaga gggtaattgga gtgcctaaaa aaagctctaa	2160
aaatagcaaa tcagtgcagc gacccctctc tacaagtgc gctttttata gaaattctga	2220
acagatatat ctatTTTTTat gaaaaggaaa atgatgcggt aacaattcag gttttaaac	2280
agcttatcca aaagattcga gaagacctcc cgaatcttga atccagtga gaaacagagc	2340
agattaacaa acattttcat aacacactgg agcatttgcg cttgcggcgg gaatcaccag	2400
aatccgaggg gccaatTTTat gaaggtctca tcctttaaaa aggaaatagc tcaccatact	2460
cctttccatg tacatccagt gagggTTTTa ttacgctagg tttcccttcc atagattgtg	2520
cctttcagaa atgctgaggt aggtttccca tttcttacct gtgatgtgtt ttaccagca	2580
cctccggaca ctcaccttca ggaccttaat aaaattattc acttggttaag tgttcaagtc	2640
tttctgatca cccaagtag catgactgat ctgcaattta aaattcctgt gatctgtaaa	2700
aaaaaaaaa aaaaaaaaaa caaaaccac aagcattat cttggctact aatgaagctc	2760
tcctTTTTTt tgtttgTTTg tttgcttcat tgttgattgt gtattttctt cattcctggg	2820
gagtactaac ccaaagcgt ctgtctcttg ttttctagtc cagtttgaga ttaatttaga	2880
agaaaggaat actgtatgtg aaattcatct tgggctttcc cctaaattgc aagataaggc	2940
catgtgtaag attttcccta aaactagaat atattaatgc atgtttgaga attttaaagc	3000
accatggtca aaaccagaag ctatatTTTTg catatttgga ctcagccatc cattaagaac	3060
ccatgttgtc ctctggacat atttatcaat ataattgggt tttaaatagt ataaaagaaa	3120
acttgtgac tatataattt atgtatcacc ttcattgtaa atttagcagg aaatgcatca	3180
caattatgat tttttTTTTt gcaccagtga aacaataaag atgctattaa caaaaaaaaa	3240
aaaaaa	3246

<210> 81

<211> 3182

<212> DNA

<213> NM\_017583.3| Homo sapiens tripartite motif-containing 44 (TRIM44), mRNA

<400> 81

ggaggctgag cgggcggcgc gacgcggggg ccgacggggg cgccgggtgg ccgcgccgga 60

agtgccttgc gcggcagagg aagcgcaggg acagagcgga gcaggccgag ccggcggaaa	120
gggtctttgc tgctgcgcc gggcaggggc tgccgcggcc ccaggtccc cttcgagacg	180
cggcgcggtc caggcgggag gcgactccct aggaagggac ccggggcggg aggaggaagt	240
gaggccgcgc ggaaggaagg cggcgagccc cggggcccg aggccttggc cgcgtcacag	300
caccacatg gcctctggag tgggcgcggc cttcgaggaa ctgcctcacg acggcacgtg	360
tgacgagtgc gagcccagc aggtccggg ggccgaggaa gtgtgccgag aatgcggctt	420
ctgctactgc cgccgccatg ccgaggcgca caggcagaag ttcctcagtc accatctggc	480
cgaatacgtc cacggctccc aggcctggac cccgccagct gacggagagg gggcggggaa	540
ggaagaagcg gaggtcaagg tggagcagga gagggagata gaaagcgagg caggggaaga	600
gagtgagtcg gaggaagaga gcgagtcaga ggaagagagc gagacagagg aagagagtga	660
ggatgagagc gatgaggaga gtgaagaaga cagcgaggaa gaaatggagg atgagcaaga	720
aagcgaggcc gaagaagaca accaagaaga aggggaatcc gaggcggagg gagaaactga	780
ggcagaaaagt gaatttgacc cagaaataga aatggaagca gagagagtgg ccaagaggaa	840
gtgtccggac catgggcttg atttgagtac ctattgccag gaagataggc agctcatctg	900
tgtcctgtgt ccagtcattg gggctacca gggccaccaa ctctccacc tagacgaagc	960
ctttgaagaa ttaagaagca aagactcagg tggactgaag gccgctatga tcgaattgggt	1020
ggaaaggttg aagttcaaga gctcagaccc taaagtaact cgggaccaa tgaagatggt	1080
tatacagcag gaatttaaga aagttcagaa agtgattgct gatgaggagc agaaggccct	1140
tcatctagtg gacatccaag aggcaatggc cacagctcat gtgactgaga tactggcaga	1200
catccaatcc cacatggata ggttgatgac tcagatggcc caagccaagg aacaacttga	1260
tacctctaat gaatcagctg agccaaaggc agagggcgat gaggaaggac ccagtgggtg	1320
cagtgaagaa gaggacacat gaaggcttgc tacccccagt ggaaaatcat cccctccctt	1380
tgtgtgtatg tgacagcgtg tatgtaacgg cttctgattt ctgtgaaagc tgctcagcaa	1440
caaacgtact tccaccagat gtgtccccag atccacagca ggcacatatc tctccaagg	1500
atgaccagtt ttatgcttac tgtgtgcttc tcatcccctg gttgtggtag gtcaaggaaa	1560
agagccccct tgatccacca ggagcaatta agaaagggtcc ttcaggtaat ccctcaatgg	1620
ctgctttgaa cttactcagg aaagccagcc ccataatat tgtattacca aacagtatcg	1680
ctttgttagg aaggatctgg aataatcttg aaggggaagc agagttttct ccctgcctat	1740
taacaaaaac ccaattttgt tcatattgaa gcatgaaata aatgagagca aggtagggcc	1800
aaattaactc ttgtggacag tccctaaaag tccagttcta catttgtaga aattgtggtg	1860
ccatgaatta agatggatga ctggaaaaag gtgttgagaa aagagttaaa gatgaggaag	1920
agatattttt agtatatgaa gttatccagg acttgatatt cataattcag tgctgtggaa	1980
atgaaaaaaa tgattgaaga ggtggaacgg aatgacctt agggggaaaa aaaaggacca	2040
aagaagtctg attaaaagtt gaaatcagta tttctgaatt caaattgctt gaatttccaa	2100

aatagtcagt aaaggatcta atagaaccag aattatttgg gtgaattctg caggttttat	2160
gggcttgtca caacgtgaag ggctggaatg tatattacca aatgggaatt tccattgtag	2220
gtttttgcta gtcccacccc catttttagcc taatttggct taaacgcagt atggggagaa	2280
ttgttcccat tccatgtgtt ctgaattcag ctcatctccc agcatataga tatatcctcc	2340
tttaactccg accagaaccc ttcttctgtt ggcactcccc acccatagac cttcagatca	2400
tctcccacac cctggatctc actctcctct tagtaacaga gacactcctg aggttggact	2460
tccttgcttt tctctacttc caaatcaca tttcttaca ccaagctttg tgctcccgag	2520
taagcagggg tgtactaggg gaatgtaaaa ctgcaaacctt aaaaacctgc atcttcttga	2580
agcatcagtt ttacttacca aatggtttag agtcataaga tgacctattt ttatataaaa	2640
gttatattat agaataaaat gttcatacgc atagactgtt aagataaaaa aataggaatc	2700
ttgcaaggta attcttattt gcaagtgggt tatgtgttca ctctcctcta cctttatggt	2760
attttggtgt tcacttacga agcatacaac tagaaccata tccaagcaga ctctgggttg	2820
ctgttaaccc agggcctaga cttctagtgc ctctgaggca gaaccaaagg agcctgcact	2880
ggggaaaatc ccttttctg cctgcctgtc tgcctgtgac ctgtgtacgt attacaggct	2940
ttaggaccag ctgattgtta tgcttgagg atggttttga aacagaaaca atacttgttt	3000
actgtaggaa tcctatttat attatttttc agtcctgtga atgctgtgaa aagatttatt	3060
cctttgaggc caggaagctc ccaggcatat atgcttctag gttaggattg tcctgactca	3120
ctaaagatgc caggatattg gggctgaggg gagtttgagg tgttaaaaaa aaaaaaaaaa	3180
aa	3182

<210> 82

<211> 4930

<212> DNA

<213> NM\_020182.3| Homo sapiens transmembrane, prostate androgen induced RNA (TMEPAI), transcript variant 1, mRNA

<400> 82	
aaacccgata tccttggact tgaatgagga ggaggaggcg gcggcggcgg cggcggcgga	60
ggcgctcggc tggggaaagc tagcggcaga ggctcagccc cggcggcagc gcgcgccccg	120
ctgccagccc attttccgga cgccacccgc gggcactgcc gacgcccccg gggctgccga	180
ggggaggccg ggggggcgca gcggagcgcg gtcccgcgca ctgagccccg cggcgccccg	240
ggaacttggc ggcgaccga gccggcgag ccggggcgcg cctccccgc gcgcgcctc	300
ctgcatgcgg ggccccagct ccgggcgccc gccggagccc cccccggccg cccccgagcc	360
ccccgcgccc cgcgccgcgc cgccgcgccc tccatgcacc gcttgatggg ggtcaacagc	420
accgcccggc ccgcccggcg gcagcccaat gtctcctgca cgtgcaactg caaacgctct	480

ttgttccaga gcatggagat cacggagctg gagtttgttc agatcatcat catcgtggtg	540
gtgatgatgg tgatggtggt ggtgatcacg tgcctgctga gccactacaa gctgtctgca	600
cggtccttca tcagccggca cagccagggg cggaggagag aagatgccct gtcctcagaa	660
ggatgcctgt ggccctcgga gagcacagtg tcaggcaacg gaatcccaga gccgcaggtc	720
tacgccccgc ctccggccac cgaccgcctg gccgtgccgc ctttcgcca gcgggagcgc	780
ttccaccgct tccagccac ctatccgtac ctgcagcacg agatcgacct gccaccacc	840
atctcgctgt cagacgggga ggagcccca ccctaccagg gccctgcac cctccagctt	900
cgggaccccg agcagcagct ggaactgaac cgggagtcgg tgcgcgcacc cccaacaga	960
accatcttcg acagtgcct gatggatagt gccaggctgg gcggccctg cccccccagc	1020
agtaactcgg gcatcagcgc cacgtgctac ggcagcggcg ggcgcagga ggggccgcg	1080
cccacctaca gcgaggtcat cggccactac ccggggctct ctttcagca ccagcagagc	1140
agtgggccgc cctccttgct ggaggggacc cggctccacc acacacacat cgcgcccta	1200
gagagcgcag ccatctggag caaagagaag gataaacaga aaggacacc tctctaggt	1260
ccccaggggg gccgggctgg ggctgcgtag gtgaaaaggc agaacactcc gcgcttcta	1320
gaagaggagt gagaggaagg cggggggcgc agcaacgcct cgtgtggccc tcccctcca	1380
cctccctgtg tataaatatt tacatgtgat gtctggtctg aatgcacaag ctaagagagc	1440
ttgcaaaaaa aaaaagaaaa aagaaaaaa aaaaccacgt ttctttgttg agctgtgtct	1500
tgaaggcaaa agaaaaaaa tttctacagt agtctttctt gtttctagtt gagctgcgtg	1560
cgtgaatgct tattttcttt tgtttatgat aatttcactt aactttaaag acatatttgc	1620
acaaaacctt tgtttaaaga tctgcaatat tatatatata aatatatata agataagaga	1680
aactgtatgt gcgagggcag gagtattttt gtattagaag aggcctatta aaaaaaaag	1740
ttgttttctg aactagaaga ggaaaaaat ggcaattttt gagtgccaag tcagaaagtg	1800
tgtattacct tgtaaagaaa aaaattacaa agcagggggt tagagttatt tatataaatg	1860
ttgagatttt gcaacttttt ttaatatata tatgtcagtg cttgcttgat ggaaacttct	1920
cttgtgtctg ttgagacttt aaggagagaa tgcggaatt tcagagtcgc ctgacggcag	1980
aggggtgagcc cccgtggagt ctgcagagag gccttgacca ggagcggcgg gctttcccga	2040
ggggccactg tccctgcaga gtggatgctt ctgcctagtg acagggtatc accacgttat	2100
atattcccta ccgaaggaga caccttttcc cccctgacct agaacagcct ttaaatcaca	2160
agcaaaatag gaaagttaac cacggaggca ccgagttcca ggtagtggtt ttgcctttcc	2220
caaaaatgaa aataaactgt taccgaagga attagttttt cctcttcttt ttccaactg	2280
tgaagggtccc cgtgggggtg agcatggtgc ccctcacaag ccgcagcggc tgggtcccgg	2340
gctaccaggg acatgccaga gggctcgatg acttgtctct gcagggcgct ttggtggttg	2400
ttcagctggc taaagggttca ccggtgaagg caggtgcggt aactgccgca ctggacccta	2460

ggaagcccca ggtattcgca atctgacctc ctctgtctg tttcccttca cggatcaatt	2520
ctcacttaag aggccaataa acaacccaac atgaaaaggt gacaagcctg ggtttctccc	2580
aggataggtg aaaggggttaa aatgagtaaa gcagttgagc aaacaccaac ccgagcttcg	2640
ggcgcagaat tcttcacctt ctcttccccct ttccatctcc tttccccgcg gaaacaacgc	2700
ttcccttctg gtgtgtctgt tgatctgtgt tttcatttac atctctctta gactccgctc	2760
ttgttctcca ggttttcacc agatagattt ggggttggcg ggacctgctg gtgacgtgca	2820
ggtgaaggac aggaaggggc atgtgagcgt aaatagaggt gaccagagga gagcatgagg	2880
ggtggggctt tgggaccac cggggccagt ggctggagct tgacgtcttt cctccccatg	2940
ggggtgggag ggcccccagc tggaagagca gactcccagc tgctaccccc tcccttccca	3000
tgggagtggc tttccatttt gggcagaatg ctgactagta gactaacata aaagatataa	3060
aaggcaataa ctattgtttg tgagcaactt tttataact tccaaaacaa aaacctgagc	3120
acagttttga agttctagcc actcgagctc atgcatgtga aacgtgtgct ttacgaaggt	3180
ggcagctgac agacgtgggc tctgcatgcc gccagcctag tagaaagttc tcgttcattg	3240
gcaacagcag aacctgcctc tccgtgaagt cgtcagccta aaatttgttt ctctcttgaa	3300
gaggattctt tgaaaaggtc ctgcagagaa atcagtacag gttatcccga aaggtacaag	3360
gacgcacttg taaagatgat taaaacgtat ctttccctta tgtgacgcgt ctctagtgcc	3420
ttactgaaga agcagtgaca ctcccgtcgc tcggtgagga cgttcccgga cagtgcctca	3480
ctcacctggg actggtatcc cctcccaggg tccaccaagg gctcctgctt ttcagacacc	3540
ccatcatcct cgcgcgtcct caccctgtct ctaccagga ggtgcctagc ttggtgaggt	3600
tactcctgct cctccaacct ttttttgcca aggtttgtac acgactcca tctaggctga	3660
aaacctagaa gtggaccttg tgtgtgtgca tgggtgcagc ccaaagccag gctgagacag	3720
tcctcatatc ctcttgagcc aaactgtttg ggtctcgttg cttcatggta tggctctggat	3780
ttgtgggaat ggctttgcgt gagaaagggg aggagagtgg ttgctgccct cagccggctt	3840
gaggacagag cctgtccctc tcatgacaac tcagtgttga agcccagtggt cctcagcttc	3900
atgtccagtg gatggcagaa gttcatgggg tagtggcctc tcaaaggctg ggcgcattcc	3960
aagacagcca gcaggttgtc tctggaaacg accagagtta agctctcggc ttctctgctg	4020
aggggtgcacc ctttccctta gatggtagtt gtcacgttat ctttgaaaac tcttgactg	4080
ctcctgagga ggccctcttt tccagtagga agttagatgg gggttctcag aagtggctga	4140
ttggaagggg acaagcttcg tttcaggggt ctgccgttcc atcctgggtc agagaaggcc	4200
gagcgtggct ttctctagcc ttgtcactgt ctccctgcct gtcaatcacc accttccctc	4260
cagaggagga aaattatctc ccctgcaaag cccggttcta cacagatttc acaaattgtg	4320
ctaagaaccg tccgtgttct cagaaagccc agtgtttttg caaagaatga aaagggaccc	4380
catatgtagc aaaaatcagg gctgggggag agccgggttc attccctgtc ctcatgggtc	4440
gtccctatga attgtacgtt tcagagaaat ttttttccct atgtgcaaca cgaagcttcc	4500

```

agaaccataa aatatcccgt cgataaggaa agaaaatgtc gttgttggtg tttttctgga 4560
aactgcttga aatcttgctg tactatagag ctgagaagga cacagcccgt cctcccctgc 4620
ctgcctgatt ccatggctgt tgtgctgatt ccaatgcttt cacgttggtt cctggcgtgg 4680
gaactgctct cctttgcagc cccatttccc aagctctggt caagttaaag ttatgtaagc 4740
tttccgtggc atgcggggcg cgcaccacg tccccgctgc gtaagactct gtatttggat 4800
gccaatccac aggcctgaag aaactgcttg ttgtgtatca gtaatcatta gtggcaatga 4860
tgacattctg aaaagctgca atacttatac aataaatttt acaattcttt ggaaaaaaaa 4920
aaaaaaaaa 4930

```

<210> 83

<211> 702

<212> DNA

<213> NM\_014183.2| Homo sapiens dynein, cytoplasmic, light polypeptide 2A (DNCL2A), transcript variant 1, mRNA

```

<400> 83
cgcagaaagg cacaggactc gctaagtgtt cgctacgcgg ggctaccgga tcggtcggaa 60
atggcagagg tggaggagac actgaagcga ctgcagagcc agaaggaggt gcagggaatc 120
atcgtcgtga acacagaagg cattcccatc aagagcacca tggacaaccc caccaccacc 180
cagtatgcca gcctcatgca cagcttcac ctgaaggcac ggagcaccgt gcgtgacatc 240
gacccccaga acgatctcac cttccttcga attcgttcca agaaaaatga aattatgggt 300
gcaccagata aagactatctt cctgattgtg attcagaatc caaccgaata agccactctc 360
ttggctccct gtgtcattcc ttaatttaat gcccccaag aatgttaatg tcaatcatgt 420
cagtggacta gcacatggca gtcgcttgga acccactcac accaatccag tgaccgtgtg 480
tgggctggcg gctcttctcc cccaccaacg gaaccctgt gtgcaccaac cttccccaga 540
gctccggagc gccctctcct cacttccagg ttttgagca agagcttgca ggaagcccg 600
accagcttc cttctgacct tcagttcact ttgtcgccct tggagaaagc tgtttttctt 660
taactaaaaa taaccaaata gcttaaaaaa aaaaaaaaaa aa 702

```

<210> 84

<211> 2100

<212> DNA

<213> NM\_015907.2| Homo sapiens leucine aminopeptidase 3 (LAP3), mRNA

```

<400> 84
ctgccccatcc gtcccgcccc ctagacgcac gtccgctcgc ccggcgcccc agccagtcgg 60

```



cgcgacgccc	gtctgcgccc	cgaaagcccc	gcccccaaggc	gcgcccgc	accgctctcc	120
acgtgctcgc	tggagggcgg	tgcgaggggc	cgagccgaca	agatgttctt	gctgcctctt	180
ccggctgcgg	ggcgagtagt	cgtccgacgt	ctggccgtga	gacgtttcgg	gagccggagt	240
ctctccaccg	cagacatgac	gaagggcctt	gttttaggaa	tctattccaa	agaaaaagaa	300
gatgatgtgc	cacagttcac	aagtgcagga	gagaattttg	ataaattggt	agctggaaag	360
ctgagagaga	ctttgaacat	atctggacca	cctctgaagg	caggggaagac	tcgaaccttt	420
tatggtctgc	atcaggactt	ccccagcgtg	gtgctagttg	gcctcggcaa	aaaggcagct	480
ggaatcgacg	aacaggaaaa	ctggcatgaa	ggcaaagaaa	acatcagagc	tgctgttgca	540
gcggggtgca	ggcagattca	agacctggag	ctctcgtctg	tggaggtgga	tccctgtgga	600
gacgctcagg	ctgctgcgga	gggagcgggtg	cttgggtctct	atgaatacga	tgacctaaag	660
caaaaaaaga	agatggctgt	gtcggcaaaag	ctctatggaa	gtggggatca	ggaggcctgg	720
cagaaaggag	tcctgtttgc	ttctgggcag	aacttggcac	gccaattgat	ggagacgcca	780
gccaatgaga	tgacgccaac	cagatttgct	gaaattattg	agaagaatct	caaaagtgtc	840
agtagtaaaa	ccgaggtcca	tatcagaccc	aagtcttgga	ttgaggaaca	ggcaatggga	900
tcattcctca	gtgtggccaa	aggatctgac	gagccccag	tcttcttgga	aattcactac	960
aaaggcagcc	ccaatgcaaa	cgaaccaccc	ctggtgtttg	ttgggaaagg	aattaccttt	1020
gacagtgggtg	gtatctccat	caaggcttct	gcaaatatgg	acctcatgag	ggctgacatg	1080
ggaggagctg	caactatatg	ctcagccatc	gtgtctgctg	caaagcttaa	tttgcccatt	1140
aatattatag	gtctggcccc	tctttgtgaa	aatatgcca	gcggcaaggc	caacaagccg	1200
ggggatgttg	ttagagccaa	aaacgggaag	accatccagg	ttgataaacac	tgatgctgag	1260
gggaggctca	tactggctga	tgcgctctgt	tacgcacaca	cgtttaaccc	gaaggctcatc	1320
ctcaatgccg	ccaccttaac	aggtgccatg	gatgtagctt	tgggatcagg	tgccactggg	1380
gtctttacca	attcatcctg	gctctggaac	aaactcttcg	aggccagcat	tgaaacaggg	1440
gaccgtgtct	ggaggatgcc	tctcttcgaa	cattatacaa	gacaggttgt	agattgccag	1500
cttgctgatg	ttaacaacat	tggaaaatac	agatctgcag	gagcatgtac	agctgcagca	1560
ttcctgaaaag	aattcgtaac	tcatcctaag	tgggcacatt	tagacatagc	aggcgtgatg	1620
accaacaaag	atgaagttcc	ctatctacgg	aaaggcatga	ctgggaggcc	cacaaggact	1680
ctcattgagt	tcttacttcg	tttcagtcaa	gacaatgctt	agttcagata	ctcaaaaatg	1740
tcttcactct	gtcttaaatt	ggacagttga	acttaaaagg	tttttgaata	aatggatgaa	1800
aatcttttaa	cggagacaaa	ggatggtatt	taaaaatgta	gaacacaatg	aaatttgtat	1860
gccttgatgt	ttttttcatt	tcacacaaag	atttataaag	gtaaagttaa	tatcttactt	1920
gataaggatt	tttaagatac	tctataaatg	attaaaattt	ttagaacttc	ctaatactt	1980
ttcagagtat	atgtttttca	ttgagaagca	aaattgtaac	tcagatttgt	gatgctagga	2040

acatgagcaa actgaaaatt actatgcact tgtcagaaac aataaatgca acttggtgtg 2100

<210> 85

<211> 1510

<212> DNA

<213> NM\_018478.1| Homo sapiens chromosome 20 open reading frame 35  
(C20orf35), mRNA

<400> 85  
cgagtgtggc caaggggtgcc ggaggcaggg ttcgggtgcg tagtcgttgc gtgggcgctg 60  
cccaaaaggc gcagagcatc aagtgtgctg gggcagaacc ggcgcgggcg cccgccgcgg 120  
gtctgcgcgg ggcggggggcg cagcaagtgc atccgagcga gcggagacta gcgcaccggc 180  
gtcgggtggcg aggggtggtgc agaggagtcc ggctggggcg agggaggaag gatgggtgcg 240  
ggtaactttt tgaccgcctt ggaagtacca gtagccgcgc tcgcaggggc tgcctccgac 300  
cgccgggcga gctgcgagcg agtgagcccg ccaccgcccc tccccactt ccgcctcggc 360  
acgaggcctc ttcctcgttc ccggctccca gggcccgtgt ccaggccgga gccagggggc 420  
ccactgttgg gatgctggct gcagtggggc gcccacagcc cagggtcccct ctgtcttctc 480  
tttcgacttt gcagctgtac ttgttttgct cctctaccg caggagctga catggacca 540  
aatcctcggg ccgccctgga gcgccagcag ctccgccttc gggagcggca aaaattcttc 600  
gaggacattt tacagccaga gacagagttt gtctttcctc tgtcccatct gcatctcgag 660  
tcgcagagac ccccatagg tagtatctca tccatggaag tgaatgtgga cactctggag 720  
caagtagaac ttattgacct tggggacccg gatgcagcag atgtgttctt gccttgcgaa 780  
gatcctccac caacccccca gtcgtctggg gtggacaacc atttggagga gctgagcctg 840  
ccggtggcta catcagacag gaccacatct aggacctcct cctcctcctc ctccgactcc 900  
tccaccaacc tgcatagccc aaatccaagt gatgatggag cagatacgcc cttggcacag 960  
tcggatgaag aggaggaaag gggatgatga ggggcagagc ctggagcctg cagctagcag 1020  
tgggcccctg cctacagact gaccacgctg gctattctcc acatgagacc acaggcccag 1080  
ccagagcctg tcgggagaag accagactct ttacttgag taggcaccag aggtgggaag 1140  
gatggtggga ttgtgtacct ttctaagaat taaccctctc ctgctttact gctaattttt 1200  
tcctgctgca accctcccac cagtttttgg cttactcctg agatatgatt tgcaaatgag 1260  
gagagagaag atgaggttgg acaagatgcc actgcttttc ttagcactct tccctcccct 1320  
aaaccatccc gtagtcttct aatacagtct ctcagacaag tgtctctaga tggatgtgaa 1380  
ctccttaact catcaagtaa ggtggtactc aagccatgct gcctccttac atcctttttg 1440  
gaacagagca cggtataaat aataaactaa taataatatg ccaacaaaaa aaaaaaaaaa 1500  
aaaaaaaaa 1510

<210> 86

<211> 3105

<212> DNA

<213> NM\_030674.2| Homo sapiens solute carrier family 38, member 1 (SLC38A1), mRNA

<400> 86  
gcacgaggga ctggggcggc cacgcactcc gccagaaggt cgccaggagc ctccgccctt 60  
caccttcctc ggaaatccgc caggccacgc aagctccctg cccaaccctt actgacgggg 120  
gccacatttt cccggcctcc gcagccagac cttgacacaa aggacatcaa actgccgagg 180  
gtaaaaaccc cggaagggcg gacacctcca catcgctttt tgccaccttt ccttttatit 240  
ccggagatat ttattgagtg tctactgtgt gccaggcact atatctatgt gcatagaaaa 300  
accctggaag gccatacaac aatatatata gagtgatcgt ctctgcttgc tgagctaaca 360  
ggggtgtcaa gcttccatit tggtatctac ttctaaatac actcagaaca ggagaaatit 420  
ggactaatit tcaaaactaca gacactttct aatcatgatg catttcaaaa gtggactcga 480  
attaactgag ttgcaaaaca tgacagtgcc cgaggatgat aacattagca atgactccaa 540  
tgatttcacc gaagtagaaa atggtcagat aaatagcaag tttatttctg atcgtgaaa 600  
tagaagaagt ctcaaaaaca gccatttgga aaaaaagaag tgtgatgagt atattccagg 660  
tacaacctcc ttaggcatgt ctgtttttta cctaagcaac gccattatgg gcagtgggat 720  
tttgggactc gccittgccc tggcaaacac tggaatccta ctttttctgg tacttttgac 780  
ttcagtgaca ttgctgtcta tatattcaat aaacctccta ttgatctgtt caaaagaaac 840  
aggctgcatg gtgtatgaaa agctggggga acaagctttt ggaccacag ggaagtctgt 900  
aatctttgga gccacctctc tacagaacac tggagcaatg ctgagctacc tcttcacgt 960  
aaaaaatgaa ctacctctg ccataaagtt tctaattgga aaggaagaga cattttcagc 1020  
ctggtacgtg gatggccgcg ttctgggtgt gatagttacc tttggcataa ttctccctct 1080  
gtgtctcttg aagaacttag ggtatcttgg ctatactagt ggattttcct tgagctgtat 1140  
ggtttttttc ctaattgtgg ttattttaca gaaatttcaa attccctgca ttgttccaga 1200  
gctaaattca acaataagtg ctaattcaac aaatgctgac acgtgtacgc caaaatatgt 1260  
taccttcaat tcaaagaccg tgtatgcttt acccaccatt gcatttgcat ttgtttgcca 1320  
cccgtcagtc ctgccaatit acagtgagct taaagaccga tcacagaaaa aaatgcagat 1380  
ggtttcaaac atctcctttt tcgccatgtt tgttatgtac ttcttgactg ccatttttgg 1440  
ctacttgaca ttctatgaca acgtgcagtc cgacctcctt caciaaatatc agagtaaaga 1500  
tgacattctc atcctgacag tgccgctggc tgtcattgtt gctgtgatcc tcacagtgcc 1560  
ggtgttattt ttcacggatc gttcatcttt atttgaactg gctaagaaaa caaagtttta 1620

tttatgtcgt	cataccgtgg	ttacctgcat	actcttggtt	gttatcaact	tgttggtgat	1680
cttcataccc	tccatgaagg	atatttttgg	agtcgtagga	gttacatctg	ctaacatgct	1740
tatttttcatt	cttccttcat	ctctttatit	aaaaatcaca	accaggatgg	agataaagga	1800
actcaaagaa	tttgggctgc	ccttttcttg	ggcctggggg	tgttgttctc	cttggtcagc	1860
attcccttgg	tcatctatga	ctgggcctgc	tcatcgagta	atggtgaagg	ccactgaaac	1920
ccgccgagaa	aaagaaacat	ccctgttgte	tgctcagtca	agtccccaca	catcagcaat	1980
ctctcaccac	ttcttttgca	agtttacaga	agcaaacaga	aatgtacagg	atacttaaaa	2040
tggaataact	ttttggttgc	aaaacagaga	catggttcta	taatgcttca	tgtccctcca	2100
agatttgaga	tcaatttagg	gattgtgaaa	tttttttttc	aaatttcata	caatcatatt	2160
tcccagtact	tttcacaatc	attttttacc	catctaactc	tatgttttgt	ggcttcccgg	2220
tctcttagaa	ctttgaaaac	atgatataca	ataatgttta	tttattatac	atccagattc	2280
tgaaataatt	ttcctactga	tgttcagctc	acactatctg	taccttttta	gaagagaaaa	2340
gaatcttgaa	ttgtatatat	ttattttgct	ttacagaaaa	aaatggtttc	gtaaataatt	2400
tgccatattt	gggtaacata	gcacatggag	ataatcatct	gaaagttata	gggcactgcc	2460
actgctgaat	cagagcatgc	ccaatatttg	agggtggctct	gatttcctgg	cagctgaact	2520
cgggtagtcc	agtggcctag	ctggtaccac	atctattccc	atccagagac	attctctggc	2580
aagtgttctc	agctgaaaag	tggttgggga	tgatttcttac	cttggttaatt	aaatgaagct	2640
acacatttgg	gtaatctagc	aaatgaagta	ttttttccct	cttggtcaact	tgtgtcagag	2700
ttactctggg	ctgagtcaac	tttcgctggg	gaaaacctat	ggaacctact	gcaaaaagat	2760
tgtccaaaat	gcctaagaaa	atactcctct	gatgcattta	gccttcaacc	ctacctgtct	2820
tgctgaaggg	agaaaaatgt	tttagtacat	tataggccca	gcagctttta	ttcatgtcca	2880
ccagctagtt	gcacagagaa	tcatgtgtac	ctaactaagg	atgatctagg	ataagtaact	2940
cctgttttat	attgagtatt	ttagggaagt	ctttaaaga	cttggtttat	atctataaat	3000
ctaggttatt	acaaatacaa	gaattttgta	ccttaaataa	gcctcatttc	tatttcttct	3060
tcattaattc	tccatctagt	cttgtgaaaa	aaaaaaaaaa	aaaaa		3105

<210> 87

<211> 2711

<212> DNA

<213> NM\_016028.4| Homo sapiens suppressor of variegation 4-20 homolog 1 (Drosophila) (SUV420H1), transcript variant 2, mRNA

<400> 87

ggtgctgcgg	cccgcgccgc	catcttggat	tttactctcc	atttttctct	ggaattattt	60
ttggtgatta	attttctggg	ggggactggg	acgcggggcc	cggcggcgcg	gccccgcac	120

gcagcggccg ggcagcgggg cctgggacgc gccccgagga ggagcggggc ggcgcaggcg	180
gagagaacat tgaaagtatt ctctaagcta tttgaagaga gtgactaaat gcacctgggt	240
caggctgtct gtgggtatga agtggttggg agaatccaag aacatggtgg tgaatggcag	300
gagaaatgga ggcaagtgtg ctaatgacca tcagcagaat caatcaaaat tacagcacac	360
ggggaaggac accctgaagg ctggcaaaaa tgcagtcgag aggaggtcga acagatgtaa	420
tggttaactcg ggatttgaag gacagagtcg ctatgtacca tcctctggaa tgtccgccaa	480
ggaactctgt gaaaatgatg acctagcaac cagtttgggt cttgatccct atttaggttt	540
tcaaacacac aaaatgaata ctagcgcctt tccttcgagg agctcaaggc atttttcaaa	600
atctgacagt ttttctcaca acaaccctgt gagatttagg cctattaaag gaaggcagga	660
agaactaaag gaagtaattg aacgttttaa gaaagatgaa cacttggaag aagccttcaa	720
atgtttgact tcaggcgaat gggcacggca ctattttctc aacaagaata aaatgcagga	780
gaaattattc aaagaacatg tatttattta tttgcgaatg tttgcaactg acagtggatt	840
tgaaatattg ccatgtaata gatactcatc agaacaaaat ggagccaaaa tagttgcaac	900
aaaagagtgg aaacgaaatg acaaaaataga attactggtg ggttgtattg ccgaactttc	960
agaaattgag gagaacatgc tacttagaca tggagaaaac gacttcagtg tcatgtactc	1020
cacaaggaaa aactgtgctc aactctggct gggtcctgct gcgtttataa accatgattg	1080
cagacctaat tgtaagtttg tgtcaactgg tcgagataca gcatgtgtga aggctctaag	1140
agacattgaa cctggagaag aaatttcttg ttattatgga gatgggttct ttggagaaaa	1200
taatgagttc tgcgagtgtt acacttgcca aagacggggc actggtgctt ttaaattccag	1260
agtgggactg cctgcgctg ctctgttat caatagcaaa tatggactca gagaaacaga	1320
taaacgttta aataggctta aaaagttagg tgacagcagc aaaaattcag acagtcaatc	1380
tgtcagctct aacactgatg cagataccac tcaggaaaaa aacaatgcaa gtaagtaagg	1440
gagatttgat aagcatatct tttaaaagta tttcacaca atttgcttta taaagtgtgc	1500
ttcagtagtt ttaaactttt aaatactcag agagactggg acttgtgagc tttggctgca	1560
cttcaaggct ctagacgtga tttgagtaga ggcacagtct gtatcccatc tctaacttca	1620
gtaccgtcct ctagactatt tttcttgaat accttggtaa ctggatatga gttcttcac	1680
atatgttcca aggtcatcat atgttttaaa catthtcaag gtgttagaga ctgtgatgat	1740
gtcgctaagt cctgcaagaa gacaaaagga ctgagtagaa ttaaattaga ctctatacat	1800
tccagtgcct agccagtttg ttagaaaaga tgatggactt ggggaattca tagcttctgg	1860
ccttaaggct tccacctttt cattgcttgc tgacctttt caaacgaac tgactcagtt	1920
cagcagacca ccagtaccag actcagaatt gtgatagagg agcattttga acagtgccgt	1980
attgtgacat gctgtattgg ctactccaga aagtaggagt aaagatggaa aggagaaaga	2040
agcaacctct gagattccag tgggtgtgtg gggcaagatc tgatggaaac tgaaaaagag	2100
aacgaagact aaacaaagag aaaggaaaga gaagaaaccc taaatgggca aaggaaagca	2160

catcctgttt gcggagcttt gaaatattgg aaccatttct aattgctcct gtttttctgg	2220
gtaacaccag ttttctgtag ttgccactaa agcagtagac tcttgagtct cacttgctctc	2280
tgagagagac agaagttaga aagttttgac ttggcgattc cgaaagtatg cctttgttgg	2340
cacttaaatg tccagtgaga cttcttggca ccttagagcc ctctgagata ctgattattt	2400
taggttcttc tccctacttt cagatgtttt cagcccaaca ctgggtgctc tcttccacta	2460
cagagaatcc tgaagaaaag ggaagggtgtt tcccatgatg gtgaatgtca ctgccatgaa	2520
ttcctgaatc tacctgctgc tgggagtcag agtccaagca taacccgtgt agcataaaaag	2580
cagcgctgta gccctattcc agtctttttc gttaatgtcc agagtgaaca acaagagtta	2640
gtcaatcatt aactgttgac tgttgattct cataataaat gcagcataac gacaaaaaaa	2700
aaaaaaaaa a	2711

<210> 88

<211> 2977

<212> DNA

<213> NM\_022105.2| Homo sapiens death associated transcription factor 1 (DATF1), transcript variant 1, mRNA

<400> 88	
gggagcggga gggcagggcg accggaggcc gcccctcagc acctctcgcg acagcaagag	60
agcgcgagag cgcgagccga tgaccaatga agcgcccccg cgagggggcg gggcggacgg	120
cctcccggaa gcgcggaacc tcagcttccg tacttgcgca gaactcccct cgcgggcgacc	180
acgcactacg gggtggcgcc agagtcaaaa ggcgtcggcc ctctggcaag atggctgctg	240
cggaggcggt ggagcgcgga aatctggaac cgggatggcg acgtctacac tgagtcggag	300
gcgaaggagc ttactccacg ggaacagcct ctagataatc tgagttgttg aaaatacgaa	360
gcctgttact cgtgaacagt ggctgacaac agtgttgttg tgagcctggc tgtctgcttg	420
gaccagaggg tttcgtctgc cagggttttt gggtgtattt aggatttcag ggaaaagtgt	480
ccaagctttc agtgttggag caggtatgga cgacaaaggc gaccagagca atgaggaggc	540
acctaaggcc atcaaaccga ccagcaaaga gttcaggaaa acatgggggtt ttcgaaggac	600
cactatcgcc aagcgagagg gcgcagggga cgcgagggct gaccactgg agccgccacc	660
cccacagcag cagctgggccc tgtccctgcg gcgcagtggg aggcagccca agcgactga	720
gcgcgtggag cagttcctga ccattgcgcg gcgcccgggc aggaggagca tgcctgtctc	780
cctggaggat tctggtgagc ccacgtcctg ccccgccaca gacgccgaga cagcctccga	840
gggcagcgtg gaaagcgctt ctgagaccag aagcggcccc cagtctgctt ccacagctgt	900
gaaggaacga ccagcctctt ctgaaaaggt gaaaggaggg gatgaccacg atgacacctc	960
cgatagtgc agcgatggcc tgaccttgaa agagcttcag aatcgccctc gcaggaagcg	1020

ggaacaggag	cccactgaga	ggccccctgaa	agggatccag	agtcgcctgc	ggaagaagcg	1080
ccgggaggag	ggccccgccg	agactgtggg	ctccgaggcc	agtgacactg	tggagggcgt	1140
cctgcccagt	aagcaggagc	ccgagaacga	tcaggggggt	gtgtcccagg	ctgggaaaga	1200
tgacagagag	agtaagttag	agggaaaggc	ggctcaggac	atcaaagatg	aggagcctgg	1260
agacttgggc	cgaccgaagc	ctgaatgtga	gggttacgac	cccaacgccc	tgtattgcat	1320
ttgccgccag	cctcacaaca	acagggtttat	gatttgcgtg	gaccgctgtg	aagaatgggt	1380
tcatggcgat	tgtgtgggca	tttctgaggc	tcgagggagg	cttttggaaa	ggaatgggga	1440
agactatata	tgcccaaact	gcaccattct	gcaagtgcag	gatgagactc	attcagaaac	1500
ggcagatcag	caggaagcta	aatggagacc	tggagatgct	gatggcaccg	attgtacaag	1560
tataggaaca	atagagcaga	agtctagcga	agaccaaggg	ataaagggtg	gaattgagaa	1620
agctgcaa	ccaagtggca	agaagaaact	caagatcttc	cagcctgtga	tagaggcgcc	1680
tgggtgcctca	aaatgtattg	gccccgggtg	ctgtcacgtg	gcgcagcccc	actcgggtgta	1740
ctgcagtaat	gactgtatcc	tcaaacacgc	cgcagcgaca	atgaagtttc	taagctcagg	1800
taaagaacag	aagccaaagc	ctaaagaaaa	gatgaagatg	aagccagaga	agcccagtct	1860
tccgaaatgc	ggtgctcagg	caggtattaa	aatctcttct	gtgcacaaga	gaccagctcc	1920
agaaaaaaaa	gagaccacag	tgaagaaggc	agtggtggtc	cctgcgcgga	gtgaagcact	1980
cgggaaggaa	gcagcttggtg	agagcagcac	gccgtcgtgg	gcgagcgatc	acaattacaa	2040
tgcagtaaag	ccagaaaaga	ctgctgctcc	ctcgccgtca	ctgttggtata	aatgtatgta	2100
tcacctaggg	gttggcctcc	tggacccctc	ccgttcttct	tggatagcca	tcccctgggc	2160
ctgtccagga	ctgggagttg	cagctttgtg	ttaagctgat	cacagacacc	ggctgcacca	2220
tcagcgggaa	gcagagccca	tgtccaggat	gcctcctgct	gccctgtgtc	catccctagt	2280
ctgtcaggac	ttcctgtcac	tgttttccaa	agctgtaa	ctcactgggtg	aacgttcacc	2340
ttaatgattg	attctttaat	ctctgttttc	actctcaggc	tctggtaagt	attcgtattc	2400
tcttcatccc	agtctgattg	catagccaca	ctgcccggca	cgccacatcc	accctgtct	2460
gcacatgagt	tgttctgaca	acagcgtgt	atacgttca	gtttttccac	attgtccacg	2520
gccagcacat	gaaagcatca	cttctttttt	atgttggtgg	aatctttgca	agttagtgtt	2580
gcatctgatt	ttcaggtgta	catttatitt	tgactgggca	gataggggat	tttttttttt	2640
ccatgtccga	ttcacacgct	acacaccac	atgaacacat	tcgaacttcg	aaggccacac	2700
actcctgctt	cataggcccc	acggtaagt	agttcacacc	tagaactctg	tcctgaccgc	2760
aggacgcgtg	ccttggaactt	ggtattctac	atgtgactgg	ctttcttgcc	ctcgtctctt	2820
gaatgttttag	actcttaaga	tcatatcctg	cccaaattt	caaattaatg	aatgaagat	2880
atttcaaaca	gatctttgaa	acctcagatt	ctgtgggtgca	attttaatgt	tttcttgttt	2940
ctcagttttc	tgctataaaa	ctattttcaa	ttcagtc			2977

<210> 89

<211> 1047

<212> DNA

<213> NM\_018487.2| Homo sapiens hepatocellular carcinoma-associated antigen 112 (HCA112), mRNA

```
<400> 89
cccacttctc cagccagcgc cccagccctc ccgccgcccg ctcgcaggtc ccgaggagcg      60
cagactgtgt ccctgacaat gggaacagcc gacagtgatg agatggcccc ggaggcccca      120
cagcacaccc acatcgatgt gcacatccac caggagtctg ccctggccaa gtcctgctc      180
acctgctgct ctgcgctgcg gccccgggcc acccaggcca ggggcagcag ccggctgctg      240
gtggcctcgt ggggtgatgca gatcgtgctg gggatcttga gtgcagtcct aggaggattt      300
ttctacatcc gcgactacac cctcctcgtc acctcgggag ctgccatctg gacaggggct      360
gtggctgtgc tggctggagc tgctgccttc atttacgaga aacgggggtg tacatactgg      420
gccctgctga ggactctgct aacgctggca gctttctcca cagccatcgc tgccctcaaa      480
ctttggaatg aagatttccg atatggctac tcttattaca acagtgcctg ccgcatctcc      540
agctcgagtg actggaacac tccagcccc actcagagtc cagaagaagt cagaaggcta      600
cacctatgta cctccttcat ggacatgctg aaggccttgt tcagaaccct tcaggccatg      660
ctcttggtg tctggattct gctgcttctg gcattctctga cccctctgtg gctgtactgc      720
tggagaatgt tcccaaccaa agggaaaaga gaccagaagg aaatggttga agtgagtgga      780
atctagccat gcctctcctg attattagtg cctggtgctt ctgcaccggg cgccccctgca      840
tctgactgct ggaagaagaa ccagactgag gaaaagaggc tcttcaacag cccagttat      900
cctggcccca tgaccgtggc cacagccctg ctccagcagc acttgcccat tccttacacc      960
ccttccccat cctgctccgc ttcatgtccc ctctgagta gtcattgtgat aataaactct     1020
catgttattg ttcccaggaa aaaaaaa                                1047
```

<210> 90

<211> 2785

<212> DNA

<213> NM\_014454.1| Homo sapiens sestrin 1 (SESN1), mRNA

```
<400> 90
gatccgccac catggctgaa ggagagaatg aagtgagatg ggatggactc tgcagcagag      60
attcaactac tagggagaca gcattggaaa acattaggca aaccattttg aggaaaaccg      120
agtatcttcg ttcggtgaaa gaaacacctc atcgtccatc agacgggctt tcaaataaccg      180
```



agtcttcgga tgggttgaat aagctacttg ctcactgct tatgctttct aagaggtgtc	240
ccttcaaaga tgtgagagag aaaagtgagt ttattctgaa gagcatccag gaacttggca	300
ttagaattcc tcgaccacta ggacagggac caagcagatt catcccagaa aaggagatcc	360
tccaagtggg gagtgaagac gcacagatgc atgctttatt tgcagattct tttgctgctt	420
tgggccgttt ggataacatt acgttagtga tggttttcca cccacaatat ttagaaagtt	480
tcttaaaaac tcagcactat ctactgcaaa tggatgggcc gttacccta cattatcgctc	540
actacattgg aataatggct gcggcaagac atcagtgtc ctacttagtg aacctgcatg	600
taaatgattt cttcatgtt ggtggggacc ccaagtggct caatggttta gagaatgtc	660
ctcaaaaact acagaattta ggagaactta acaaagtgtt agcccataga ctttggctta	720
ttaccaaaga acacattgag ggacttttaa aagctgaaga gcacagctgg tcccttgccg	780
aattggtaca tgcagtagtt ttactcacac actatcattc tcttgctca ttcacattcg	840
gctgtggaat cagtcagaa attcattgtg atgggtggcca cacattcaga cctccttctg	900
ttagcaacta ctgcatctgt gacattaca atggcaatca cagtgtggat gagatgccg	960
tcaactcagc agaaaatgtt tctgtaagt atcttttctt tgaggttgaa gccctcatgg	1020
aaaagatgag gcagttacag gaatgtcgag atgaagaaga ggcaagtcag gaagagatgg	1080
cttcacgttt tgaaatagaa aaaagagaga gtatgtttgt cttctcttca gatgatgaag	1140
aagttacacc agcaagagct gtatctcgtc attttgagga tactagttat ggctataaag	1200
atcttcttag acatgggatg catgttccaa catttcgtgt ccaggactat tgctgggaag	1260
atcatggta ttcttttgta aatcgcttt atccagatgt gggacagttg attgatgaaa	1320
aatttcacat tgcttacaat ctacttata atacaatggc aatgcacaaa gatgttgata	1380
cctcaatgct tagacgggca atttggaact atattcactg catgtttgga ataagatatg	1440
atgattatga ctatggtgaa attaaccagc tattggatcg tagctttaaa gtttatatca	1500
aaactgttgt ttgcactcct gaaaagggtta ccaaagaat gtatgatagc ttctggaggc	1560
agttcaagca ctctgagaag gttcatgtta atctgcttct tatagaagct aggatgcaag	1620
cagaactcct ttatgctctg agagccatta cccgctatat gacctgatgc ctttccttca	1680
ttaaagatga ttctggaatg atcagcagat atagtctaca agggggaagg tactaagccc	1740
caggaccaat ggtagacaaa ataattcaga aatccattgt gccatgattc ctttagtttc	1800
tgctatTTTT ctgtggaaaa cactgctgg cacaagcagt gactgtttgg cagcttcaag	1860
tttagagctg tgaagacagg ctgccattca cagtattttg ctttttgaca gtacaagatg	1920
ctgtgtaact gttttaatac agcaaatagt aactctccaa atcctgttgc ttttatgtta	1980
aataagataa caagaattgg agcatgcaaa gaatgggact tggataatga ctttaagcttt	2040
atatgtaaag aatttttagaa gatcttggtg ctgctattcc tgctggagga atgaatagat	2100
ggctgtttca gttaagctat tagtaataaa agtgaacatt gctactatct gagcctacat	2160
acataacttg tgtgatttca aattaaactt gcattatgtg ttaattttct tgcattcaaa	2220

aaagcataga attcctactc acacagctca gcaacaacca ttttgatggt aacagttaat	2280
ttctttcatt agtttttttaa attcaggggt ctggatatta aattaaaatg gcattcttaa	2340
agattttctt caaaaagcaa tcctaaatga aagtgtgtaa attataagaa gctggcgatc	2400
ttttgatatg ctgtttcaca ggatcctgac actggagggc agctgtcttg tgcattactt	2460
gtgttcccag caccaaagtt gtgggacatg ttgctgtaga ctgctgcgca gtcctgggtg	2520
cattcagtct ctctgcctct gcctgcctcc tgggtcccccac tttaaaggct gtgcagctcc	2580
ttaaataata aagctggaaa atatttttag tcgggttatc aaatttgatt tacaaaaacg	2640
ctaactttgt ttgaaatgca aacaggtttg aaaatatgta ttaagtactt tgtattctgg	2700
aagcgtgaat tgcttttgaa gtctgtcagt attactggta tttttaaata aagaagaatt	2760
tttctccaaa aaaaaaaaaa aaaaa	2785

<210> 91

<211> 3802

<212> DNA

<213> NM\_017763.1| Homo sapiens hypothetical protein FLJ20315 (FLJ20315), mRNA

<400> 91	
aaaaaaaaaa aacttttagag aaaggaaggg ccaaaactac gacttggctt tctgaaacgg	60
aagcataaat gttcttttcc tccatttgtc tggatctgag aacctgcatt tgggtattagc	120
tagtggaagc agtatgtatg gttgaagtgc attgctgcag ctggtagcat gagtgggtggc	180
caccagctgc agctggctgc cctctggccc tggctgctga tggctaccct gcaggcaggc	240
tttgacgca caggactggt actggcagca gcggtggagt ctgaaagatc agcagaacag	300
aaagctgtta tcagagtgat ccccttgaaa atggaccca caggaaaact gaatctcact	360
ttggaagggtg tgtttgctgg tgttgctgaa ataactccag cagaaggaaa attaatgcag	420
tcccacccac tgtacctgtg caatgccagt gatgacgaca atctggagcc tggattcatc	480
agcatcgtca agctggagag tcctcgacgg gcccccgcc cctgcctgtc actggctagc	540
aaggctcgga tggcgggtga gcgaggagcc agtgctgtcc tctttgacat cactgaggat	600
cgagctgctg ctgagcagct gcagcagccg ctggggctga cctggccagt ggtgttgatc	660
tggggtaatg acgctgagaa gctgatggag tttgtgtaca agaaccaaaa ggcccatgtg	720
aggattgagc tgaaggagcc cccggcctgg ccagattatg atgtgtggat cctaattgaca	780
gtggtgggca ccattcttgt gatcatcctg gcttcgggtg tgccgcatccg gtgccgcccc	840
cgccacagca ggccggatcc gcttcagcag agaacagcct gggccatcag ccagctggcc	900
accaggaggt accaggccag ctgcaggcag gccccgggtg agtggccaga ctcagggagc	960
agctgcagct cagcccctgt gtgtgccatc tgtctggagg agttctctga ggggcaggag	1020

ctacgggtca	tttcctgcct	ccatgagttc	catcgtaact	gtgtggaccc	ctggttacat	1080
cagcatcgga	cttgccccct	ctgcgtgttc	aacatcacag	agggagattc	attttcccag	1140
tccctgggac	cctctcgatc	ttaccaagaa	ccaggtcgaa	gactccacct	cattcgccag	1200
catcccggcc	atgcccacta	ccacctccct	gctgcctacc	tggtgggccc	ttcccggagt	1260
gcagtggctc	ggccccccacg	acctgggtccc	ttcctgccat	cccaggagcc	aggcatgggc	1320
cctcggcatc	accgcttccc	cagagctgca	catccccggg	ctccaggaga	gcagcagcgc	1380
ctggcaggag	cccagcacc	ctatgcacaa	ggctggggaa	tgagccacct	ccaatccacc	1440
tcacagcacc	ctgctgcttg	cccagtgc	ctacgccggg	ccaggccccc	tgacagcagt	1500
ggatctggag	aaagctattg	cacagaacgc	agtgggtacc	tggcagatgg	gccagccagt	1560
gactccagct	cagggccctg	tcatggctct	tccagtgact	ctgtggtcaa	ctgcacggac	1620
atcagcctac	aggggggtcca	tggcagcagt	tctactttct	gcagctccct	aagcagtgac	1680
tttgacccc	tagtgtactg	cagccctaaa	ggggatcccc	agcgagtggg	catgcagcct	1740
agtgtgacct	ctcggcctcg	ttccttgag	tcgggtggtgc	ccacagggga	aaccaggtt	1800
tccagccatg	tccactacca	ccgccaccgg	caccaccact	acaaaaagcg	gttccagtgg	1860
catggcagga	agcctggccc	agaaaccgga	gtcccccagt	ccaggcctcc	tattcctcgg	1920
acacagcccc	agccagagcc	accttctcct	gatcagcaag	tcaccggatc	caactcagca	1980
gccccttcgg	ggcggtctct	taaccacacag	tgccccaggg	ccctccctga	gccagcccc	2040
ggcccagttg	acgcctccag	catctgcccc	agtaccagca	gtctgttcaa	cttgcaaaaa	2100
tccagcctct	ctgcccagca	cccacagagg	aaaaggcg	gggggtccctc	cgagcccacc	2160
cctggctctc	ggccccagga	tgcaactgtg	caccagcctt	gccagatttt	tccccattac	2220
acccccagtg	tggcatatcc	ttgggtcccca	gaggcacacc	ccttgatctg	tggacctcca	2280
ggcctggaca	agaggctgct	accagaaacc	ccaggccctt	gttactcaaa	ttcacagcca	2340
gtgtggttgt	gcctgactcc	tcgccagccc	ctggaaccac	atccacctgg	ggaggggcct	2400
tctgaatgga	gttctgacac	cgcagagggc	aggccatgcc	cttatccgca	ctgccaggtg	2460
ctgtcggccc	agcctggctc	agaggaggaa	ctcaggagagc	tgtgtgaaca	ggctgtgtga	2520
gatgttcagg	cctagctcca	accaagagtg	tgctccagat	gtgtttgggc	cctacctggc	2580
acagagtcct	gctcctggga	aaggaaagga	ccacagcaaa	caccattctt	tttgccgtac	2640
ttcctagaag	cactggaaga	ggactgggtga	tggtaggggg	tgagaggggtg	ccgtttcctg	2700
ctccagctcc	agacctgtg	tgcaaaaaac	atctgcagtg	cagcaaatcc	atgtccagcc	2760
aggcaaccag	ctgctgcctg	tggcgtgtgt	gggctggatc	ccttgaaggc	tgagtttttg	2820
agggcagaaa	gctagctatg	ggtagccagg	tggtacaaaag	gtgctgctcc	ttctccaacc	2880
cctacttgg	ttccctcacc	ccaagcctca	tggtcatacc	agccagtggg	ttcagcagaa	2940
cgcattgacac	cttatcacct	ccctccttgg	gtgagctctg	aacaccagct	ttggccctc	3000
cacagtaagg	ctgctacatc	aggggcaacc	ctggctctat	cattttcctt	ttttgcaaaa	3060

aggaccagta gcataggtga gccctgagca ctaaaaggag gggtcctga agctttccca 3120  
 ctatagtgtg gaggttctgtc cctgaggtgg gtacagcagc cttggttcct ctgggggttg 3180  
 agaataagaa tagtggggag ggaaaaactc ctccttgaag atttcctgtc tcagagtccc 3240  
 agagaggtag aaaggaggaa tttctgctgg actttatctg ggcagaggaa ggatggaatg 3300  
 aaggtagaaa aggcagaatt acagctgagc ggggacaaca aagagttctt ctctgggaaa 3360  
 agttttgtct tagagcaagg atggaaaatg gggacaacaa aggaaaagca aagtgtgacc 3420  
 cttggggttg gacagcccag agggccagct cccagtata agccatacag gccagggacc 3480  
 cacaggagag tggattagag cacaagtctg gcctcactga gtggacaaga gctgatgggc 3540  
 ctcatcaggg tgacattcac cccagggcag cctgaccact cttggcccct caggcattat 3600  
 cccatttgga atgtgaatgt ggtggcaaag tgggcagagg accccacctg ggaacctttt 3660  
 tccctcagtt agtggggaga ctagcaccta ggtaccaca tgggtattta tatctgaacc 3720  
 agacagacgc ttgaatcagg cactatgtta agaaatatat ttatttgcta atatatttat 3780  
 ccacaaaaaa aaaaaaaaaa aa 3802

<210> 92

<211> 1236

<212> DNA

<213> NM\_017918.3| Homo sapiens hypothetical protein FLJ20647 (FLJ20647), mRNA

<400> 92  
 caggcgctga cgaggagccc ggctgagggg ggatgcgccg ctgacgcctg cgggagccgc 60  
 gcgcctgggg cgggaggatg ctccagaggg gcctctggcc gtggcgcacg cggctgtctg 120  
 cgacccttg cacttgccgc ccagcgccg cgtggccgct gccgcctccg cccaggttt 180  
 tgcgtgtgaa gctgtgtgga aatgtgaaat actaccagtc acaccattat agtaccgtgg 240  
 tgccacctga tgaaataaca gttatttata gacatggcct tcccttggtgta acacttacct 300  
 tgccatctag aaaagaacgt tgtcaattcg tagtcaaacc aatgttgtca acagttggtt 360  
 cattccttca ggacctacaa aatgaagata agggatatcaa aactgcagcc atcttcacag 420  
 cagatggcaa catgatttca gcttctacct tgatggatat ttgctaatag aatgatttta 480  
 aacttgtcat taataaaata gcatatgatg tgcagtgtcc aaagagagaa aaaccaagta 540  
 atgagcacac tgctgagatg gaacacatga aatccttggc tcacagacta ttacaatct 600  
 tgcatttaga agagtctcag aaaaagagag agcaccattt actggagaaa attgaccacc 660  
 tgaaggaaca gctgcagccc cttgaacagg tgaaagctgg aatagaagct cattcggaag 720  
 ccaaaaccag tggactcctg tgggctggat tggcactgct gtccattcag ggtggggcac 780  
 tggcctggct cacgtgggtg gtgtactcct gggatatcat ggagccagtt acataacttca 840

tcacatttgc aaattctatg gtcttttttg catactttat agtcactcga caggattata	900
cttactcagc tgttaagagt aggcaatttc ttcagttctt ccacaagaaa tcaaagcaac	960
agcactttga tgtgcagcaa tacaacaagt taaaagaaga ccttgctaag gctaaagaat	1020
ccctgaaaca ggcgcgtcat tctctctgtt tgcaaagtca agtagaagaa ctcaatgaaa	1080
agaattaatc ttacagtttt aaatgtcgtc agattttcca ttatgtattg attttgcaac	1140
ttaggatgtt tttagatccc atgggttcatt ttgattgttt aatctttgtt attaaattct	1200
tgtaaaacag aaaaaaaaaa aaaaaaaaaa aaaaaa	1236

<210> 93

<211> 2096

<212> DNA

<213> NM\_024792.1| Homo sapiens membrane protein expressed in epithelial-like lung adenocarcinoma (CT120), mRNA

<400> 93	
agccacgcgg cgccagcgag gcggccggac ccgcagcccc gatgctgctg acgctggccg	60
ggggcgcgct cttcttcccc gggctcttcg cgctctgcac ctgggcgctg cgccgctccc	120
agcccggaatg gagccgcacc gactgcgtga tgatcagcac caggctgggtt tcctcgggtg	180
acgccgtgct ggccaccggc tcggggatcg tcatcattcg ctctgcgac gacgtgatca	240
ccggcaggca ctggcttgcc cggaatatg tgtggtttct gattccatac atgatctatg	300
actcgtacgc catgtacctc tgtgaatggg gccgaaccag agaccagaac cgtgcgccct	360
ccctcactct tcgaaacttc ctaagtcgaa accgcctcat gatcacacat catgcgggtca	420
ttctctttgt ccttggtgcca gtcgcacaga ggctccgggg agacctggg gacttctttg	480
tcggctgcat cttcacggca gaactgagca ctccgtttgt gtcgctgggc agggttctga	540
ttcagctaaa gcagcagcac acccttctgt acaagggtgaa tggaatcctc acgctggcca	600
ccttcctttc ctgccggatc cttctcttcc cttcatgta ctggtcctat ggccgccagc	660
agggactaag cctgctccaa gtacccttca gcatccatt ctactgcaac gtggccaatg	720
ccttcctcgt agctcctcag atctactggg tctgtctgct gtgcaggaag gcagtccggc	780
tctttgacac tccccagcc aaaaaggatg gctaaatgct cctgggagtc aggcgcagcc	840
tcacaccagc tgcctcctcc actcagcatt ccatggacca aattgtgccc tgggtagcct	900
cagactttgg gtattgataa gccgatggat ttgagttttt ctaaagaata ttcattattac	960
ctccttcttc taacttgccc tatttgcaaa agcacttttg tagtaacaac tattgggtcc	1020
tgtcagacct ccacggacag caaagtgggt ttaatgcaag cccaaggatc cttcttaagg	1080
tcttatctca agagctctgg gaggtggaag catgggggtg gatcgggtgga ccagggtggg	1140
aagtgtctgc acatctgcct gtccctgtat cagcggctac ccaccttcca aaccactcag	1200

gacagtaccc	gtggcactgg	gccccgagaa	gcaagggatg	acttggttct	tggaagtaat	1260
gtcgtcttgt	gacattggcc	tgggacaatc	attgtgggta	ggtagttatt	gatcgtttac	1320
tagataaccc	attgggttctt	tgcctcatcc	tctcatccat	gggtcagagt	tgaattctta	1380
tgtctataga	cttccaatca	gaagtctcac	tgggtggggct	gggggtgggg	gcaggcagga	1440
ggcatggatg	ggaacctgag	taggtagtgt	ggccaagaga	tcagcacaac	ctttgcaggc	1500
tgacttgcta	agtctgacag	tgacaaactt	gtgagcttac	tgcagtcagt	cacagaggct	1560
gttctttttc	acacacccct	tcatgcccgg	ctttcccat	atccacatgc	agagggcgag	1620
ctcataaaac	tacaggggaag	cgtgaaatga	tggctttggt	agctgtttac	tgggtaaccc	1680
cactgtgaca	ctgtcctttt	catgtgatgt	ggaaacctac	ttctgtcctc	caaaccatga	1740
aatgtgtcat	ctagactgca	gagtacttga	gtgctttgcc	tcccgatatg	ccagagcttg	1800
tggtcctaaag	cccatcctg	tgtgtccgtc	ctgccattta	gccacagaag	gctgcggagt	1860
gaggcggcag	ctagcctggc	cagtggctgt	cccgtggacc	gacacctgcg	cccccttctg	1920
caagcaggat	tttctggtgc	caacactcat	tcatcattcc	cgatcaacta	ggatgaattt	1980
aagactgtgc	taccatgtgt	tctcaagtgg	tagtttaaaa	agtggatttt	taaagtgcct	2040
ttcaattgtc	tgtgaacgtc	taaaggactg	atttgtctca	aaaaaaaaaa	aaaaaa	2096

<210> 94

<211> 4372

<212> DNA

<213> NM\_014314.2| Homo sapiens DEAD (Asp-Glu-Ala-Asp) box polypeptide 58 (DDX58), mRNA

<400> 94						
tagttattaa	agttcctatg	cagctccgcc	tcgctccgg	cctcatttcc	tcggaaaatc	60
cctgctttcc	ccgctcgcca	cgccctcctc	ctacccggct	ttaaagctag	tgaggcacag	120
cctgcgggga	acgtagctag	ctgcaagcag	aggccggcat	gaccaccgag	cagcgacgca	180
gcctgcaagc	cttccaggat	tatatccgga	agaccctgga	ccctacctac	atcctgagct	240
acatggcccc	ctggtttagg	gaggaagagg	tgcagtatat	tcaggctgag	aaaaacaaca	300
agggcccaat	ggaggctgcc	acactttttc	tcaagttcct	gttggagctc	caggaggaag	360
gctggttccg	tggctttttg	gatgccctag	accatgcagg	ttattctgga	ctttatgaag	420
ccattgaaag	ttgggatttc	aaaaaaattg	aaaagttgga	ggagtataga	ttacttttaa	480
aacgtttaca	accagaattt	aaaaccagaa	ttatcccaac	cgatatcatt	tctgatctgt	540
ctgaatgttt	aattaatcag	gaatgtgaag	aaattctaca	gatttgctct	actaagggga	600
tgatggcagg	tgcagagaaa	ttgggtggaat	gccttctcag	atcagacaag	gaaaactggc	660
ccaaaacttt	gaaacttgct	ttggagaaaag	aaaggaacaa	gttcagtga	ctgtggattg	720

tagagaaagg	tataaaagat	gttgaaacag	aagatcttga	ggataagatg	gaaacttctg	780
acatacagat	tttctaccaa	gaagatccag	aatgccagaa	tcttagtgag	aattcatgtc	840
caccttcaga	agtgtctgat	acaaacttgt	acagcccatt	taaaccaaga	aattaccaat	900
tagagcttgc	tttgccctgct	atgaaaggaa	aaaacacaat	aatatgtgct	cctacagggt	960
gtggaaaaac	ctttgtttca	ctgcttatat	gtgaacatca	tcttaaaaaa	ttcccacaag	1020
gacaaaaggg	gaaagttgtc	ttttttgcga	atcagatccc	agtgtatgaa	cagcagaaat	1080
ctgtattctc	aaaatacttt	gaaagacatg	ggtatagagt	tacaggcatt	tctggagcaa	1140
cagctgagaa	tgtcccagtg	gaacagattg	ttgagaacaa	tgacatcatc	attttaactc	1200
cacagattct	tgtgaacaac	cttaaaaagg	gaacgattcc	atcactatcc	atctttactt	1260
tgatgatatt	tgatgaatgc	cacaacacta	gtaaacaaca	cccgtacaat	atgatcatgt	1320
ttaattatct	agatcagaaa	cttggaggat	cttcaggccc	actgccccag	gtcattgggc	1380
tgactgcctc	ggttgggtgt	ggggatgcc	aaaacacaga	tgaagccttg	gattatatct	1440
gcaagctgtg	tgcttctctt	gatgcgtcag	tgatagcaac	agtcaaacac	aatctggagg	1500
aactggagca	agttgtttat	aagccccaga	agtttttcag	gaaagtggaa	tcacggatta	1560
gcgacaaaat	taaatacatc	atagctcagc	tgatgaggg	cacagagagt	ctggcaaaga	1620
gaatctgcaa	agacctcgaa	aacttatctc	aaattcaaaa	tagggaattt	ggaacacaga	1680
aatatgaaca	atggattgtt	acagttcaga	aagcatgcat	ggtgttccag	atgccagaca	1740
aagatgaaga	gagcaggatt	tgtaaagccc	tgtttttata	cacttcacat	ttgcggaaat	1800
ataatgatgc	cctcattatc	agtgagcatg	cacgaatgaa	agatgctctg	gattacttga	1860
aagacttctt	cagcaatgtc	cgagcagcag	gattcgatga	gattgagcaa	gatcttactc	1920
agagatttga	agaaaagctg	caggaactag	aaagtgtttc	cagggatccc	agcaatgaga	1980
atcctaaaact	tgaagacctc	tgcttcatct	tacaagaaga	gtaccactta	aaccagaga	2040
caataacaat	tctctttgtg	aaaaccagag	cacttgtgga	cgctttaaaa	aattggattg	2100
aaggaaatcc	taaactcagt	tttctaaaac	ctggcatatt	gactggacgt	ggcaaaaaca	2160
atcagaacac	aggaatgacc	ctcccggcac	agaagtgtat	attggatgca	ttcaaagcca	2220
gtggagatca	caatattctg	attgccacct	cagttgctga	tgaaggcatt	gacattgcac	2280
agtgcaatct	tgtcatcctt	tatgagtatg	tgggcaatgt	catcaaatg	atccaaacca	2340
gaggcagagg	aagagcaaga	ggtagcaagt	gcttccttct	gactagtaat	gctgggtgtaa	2400
ttgaaaaaga	acaaataaac	atgtacaaag	aaaaaatgat	gaatgactct	attttacgcc	2460
ttcagacatg	ggacgaagca	gtatttaggg	aaaagattct	gcatatacag	actcatgaaa	2520
aattcatcag	agatagtcaa	gaaaaaccaa	aacctgtacc	tgataaggaa	aataaaaaac	2580
tgctctgcag	aaagtgcaaa	gccttggcat	gttacacagc	tgacgtaaga	gtgatagagg	2640
aatgccatta	cactgtgctt	ggagatgctt	ttaaggaatg	ctttgtgagt	agaccacatc	2700
ccaagccaaa	gcagttttca	agttttgaaa	aaagagcaaa	gatattctgt	gcccgcacaga	2760

```

actgcagcca tgactgggga atccatgtga agtacaagac atttgagatt ccagttataa 2820
aaattgaaag ttttgtggtg gaggatattg caactggagt tcagacactg tactcgaagt 2880
ggaaggactt tcattttgag aagataccat ttgatccagc agaaatgtcc aaatgatatc 2940
aggtcctcaa tcttcagcta cagggaatga gtaactttga gtggagaaga aacaaacata 3000
gtgggtataa tcatggatcg cttgtacccc tgtgaaaata tattttttta aaatatcttt 3060
agcagtttgt actatattat atatgcaaag cacaaatgag tgaatcacag cactgagtat 3120
tttgtaggcc aacagagctc atagtacttg ggaaaaatta aaaagcctca tttctagcct 3180
tctttttaga gtcaactgcc aacaaacaca cagtaatcac tctgtacaca ctgggataga 3240
tgaatgaatg gaatgttggg aatttttatc tccctttgtc tccttaacct actgtaaact 3300
ggcttttgcc cttacaatc tactgaaatt gttcttttga aggttaccag tgactctggt 3360
tgccaaatcc actgggcact tcttaacctt ctatttgacc tctgcgcatt tggccctggt 3420
gagcactctt cttgaagctc tccctgggct tctctctctt ctagttctat tctagtcttt 3480
ttttattgag tctctctctt tgctgatccc ttccaagggg tcaatatata tacatgtata 3540
tactgtacat atgtatatgt aactaatata catacatata ggtatgtata tgtaatgggt 3600
atatgtactc atgttcctgg tgtagcaacg tgtggtatgg ctacacagag aacatgagaa 3660
cataaagcca tttttatgct tactactaaa agctgtccac tgtagagttg ctgtatgtag 3720
caatgtgtat ccactctaca gtggtcagct tttagtagag agcataaaaa tgataaaata 3780
cttcttgaaa acttagttta ctatacatct tgccctatta atatgttctc ttaacgtgtg 3840
ccattgttct ctttgaccat tttcctataa tgatgttgat gttcaacacc tggactgaat 3900
gtctgttctc agatcccttg gatgttacag atgaggcagt ctgactgtcc tttctacttg 3960
aaagattaga atatgtatcc aaatggcatt cacgtgtcac ttagcaagggt ttgctgatgc 4020
ttcaaagagc ttagtttgcg gtttcctgga cgtggaaaca agtatctgag ttccctggag 4080
atcaacggga tgaggtgta cagctgcctc cctcttcatg caatctggtg agcagtgggtg 4140
caggcgggga gccagagaaa cttgccagtt atataacttc tctttggctt ttcttcatct 4200
gtaaaacaag gataaactg aactgtaagg gttagtggag agtttttaaat taaaagaatg 4260
tgtgaaaagt acatgacaca gtagttgctt gataatagtt actagtagta gtattcttac 4320
taagacccaa tacaaatgga ttatttaaac caaaaaaaaaa aaaaaaaaaa aa 4372

```

<210> 95

<211> 2163

<212> DNA

<213> NM\_015515.3| Homo sapiens keratin 23 (histone deacetylase inducible) (KRT23), transcript variant 1, mRNA



<400> 95  
ggcagatgaa atataagatt catcaaccac atttgacagc ccatggcagg tttcctgttt 60  
tccatcgtcc ctctgcaggt cacagacaca cagagcccag ccgtggcagg ctcagccggg 120  
gtccggggct gctaacaacg gctacattcc tccccaggg ccaagggaaa tcctgagcgc 180  
aggccagggt tgtttggttt tgaggtgtgc tgggatgaaa ggcaccctgg aagtgggaagg 240  
taaatagaaca atggaaaaac ttcacggcaa gattagaaag atacctgagc ccaatacccg 300  
cctgatgtcg tgggccacac ctccgggtta ccagggggaag ggaggaagca aactgtcata 360  
ttgatgtggc tctaaacaac aacagtgtgc gaaggcccag gggcactttg ggattgacca 420  
agaggaaaca caagttgcac aatgatacaa tcttggttgg acaattgtca gagaagggaa 480  
ctcccacagc aaaggccata aaaccatcca gggcagtctg gggcggctca gttctgcggt 540  
gccagggagt ggagcagagc tcagccccgt cccaaacaca gatgggacca tgaactccgg 600  
acacagcttc agccagaccc cctcggcctc cttccatggc gccggagggt gctggggccg 660  
gcccaggagc tccccaggg ctcccaccgt ccatggcgggt gcggggggag cccgcatctc 720  
cctgtccttc accacgcgga gctgcccacc ccctggagggt tcttgggggt ctggaagaag 780  
cagcccccta ctaggcggaa atgggaaggc caccatgcag aatctcaacg accgcctggc 840  
ctcctacctg gagaagggttc gcgccctgga ggaggccaac atgaagctgg aaagccgcat 900  
cctgaaatgg caccagcaga gagatcctgg cagtaagaaa gattattccc agtatgagga 960  
aaacatcaca cacctgcagg agcagatagt ggatggtaag atgaccaatg ctcagattat 1020  
tcttctcatt gacaatgcca ggatggcagt ggatgacttc aacctcaagt atgaaaatga 1080  
acactccttt aagaaagact tggaaattga agtcgagggc ctccgaagga ccttagacaa 1140  
cctgaccatt gtcacaacag acctagaaca ggagggtggaa ggaatgagga aagagctcat 1200  
tctcatgaag aagcaccatg agcaggaaat ggagaagcat catgtgccaa gtgacttcaa 1260  
tgtcaatgtg aagggtggata cagggtcccag ggaagatctg attaagggtcc tggaggatat 1320  
gagacaagaa tatgagctta taataaagaa gaagcatcga gacttggaca cttggtataa 1380  
agaacagtct gcagccatgt cccaggaggc agccagtcca gccactgtgc agagcagaca 1440  
agggtgacatc cacgaactga agcgcacatt ccaggccctg gagattgacc tgcagacaca 1500  
gtacagcacg aaatctgctt tggaaaacat gttatccgag acccagtctc ggtactcctg 1560  
caagctccag gacatgcaag agatcatctc ccactatgag gaggaactga cgcagctacg 1620  
ccatgaactg gagcggcaga acaatgaata ccaagtgtg ctgggcatca aaaccacct 1680  
ggagaaggaa atcaccacgt accgacggct cctggaggga gagagtgaag ggacacggga 1740  
agaatcaaag tcgagcatga aagtgtctgc aactccaaag atcaaggcca taaccagga 1800  
gaccatcaac ggaagattag ttctttgtca agtgaatgaa atccaaaagc acgcatgaga 1860  
ccaatgaaag tttccgcctg ttgtaaaatc tttttcccc caaggaaagt ccttgcacag 1920  
acaccagtga gtgagttcta aaagataccc ttggaattat cagactcaga aacttttatt 1980

ttttttttct gtaacagtct caccagactt ctcataatgc tcttaatata ttgcactttt	2040
ctaatacaaag tgcgagttta tgagggtaaa gctctacttt cctactgcag ccttcagatt	2100
ctcatcattt tgcacatatt ttgtagccaa taaaactccg cactagcaaa aaaaaaaaaa	2160
aaa	2163

<210> 96

<211> 2881

<212> DNA

<213> NM\_007210.2| Homo sapiens UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetyl-galactosaminyltransferase 6 (GalNAc-T6) (GALNT6), mRNA

<400> 96	
atgaggctcc tccgcagacg ccacatgccc ctgcgccctg ccatggtggg ctgcgccctt	60
gtgctcttcc tcttctctct gcatagggat gtgagcagca gagaggaggc cacagagaag	120
ccgtggctga agtccctggg gagccggaag gatcacgtcc tggacctcat gctggaggcc	180
atgaacaacc ttagagattc aatgcccagg ctccaaatca gggctccaga agcccagcag	240
actctgttct ccataaacca gtcttgctc cctgggttct ataccccagc tgaactgaag	300
cccttctggg aacggccacc acaggacccc aatgcccctg gggcagatgg aaaagcattt	360
cagaagagca agtggacccc cctggagacc caggaaaagg aagaaggcta taagaagcac	420
tgtttcaatg cctttgccag cgaccggatc tccctgcaga ggtccctggg gccagacacc	480
cgaccacctg agtgtgtgga ccagaagttc cggcgctgcc cccactggc caccaccagc	540
gtgatcattg tgttccacaa cgaagcctgg tccacactgc tgcgaacagt gtacagcgtc	600
ctacacacca cccctgccat cttgctcaag gagatcatac tgggtggatga tgccagcaca	660
gaggagcacc taaaggagaa gctggagcag tacgtgaagc agctgcagggt ggtgaggggtg	720
gtgcggcagg aggagcggaa ggggttgatc accgcccggc tgctgggggc cagcgtggca	780
caggcggagg tgctcacgtt cctggatgcc cactgtgagt gcttccacgg ctggctggag	840
cccctcctgg ctccaatcgc tgaggacaag acagtgggtg tgagcccaga catcgtcacc	900
atcgacctta atacttttga gttcgccaag cccgtccaga ggggcagagt ccatagccga	960
ggcaactttg actggagcct gaccttcggc tgggaaacac ttcctccaca tgagaagcag	1020
aggcgcaagg atgaaacata ccccatcaaa tccccgacgt ttgctgggtg cctcttctcc	1080
atccccaaagt cctactttga gcacatcggg acctatgata atcagatgga gatctgggga	1140
ggggagaacg tggaaatgtc cttccgggtg tggcagtgtg ggggccagct ggagatcatc	1200
ccctgctctg tcgtaggcca tgtgttccgg accaagagcc cccacacctt cccaagggc	1260
actagtgtca ttgctcgcaa tcaagtgcgc ctggcagagg tctggatgga cagctacaag	1320
aagattttct ataggagaaa tctgcaggca gcaaagatgg cccaagagaa atccttcggt	1380

gacatttcgg aacgactgca gctgagggaa caactgcact gtcacaactt ttcctggtac	1440
ctgcacaatg tctacccaga gatgtttgtt cctgacctga cgcccacctt ctatggtgcc	1500
atcaagaacc tcggcaccaa ccaatgcctg gatgtgggtg agaacaaccg cggggggaag	1560
cccctcatca tgtactcctg ccacggcctt ggcggcaacc agtactttga gtacacaact	1620
cagagggacc ttcgccacaa catcgcaaag cagctgtgtc tacatgtcag caaggggtgct	1680
ctgggccttg ggagctgtca cttcactggc aagaatagcc aggtcccca ggacgaggaa	1740
tgggaattgg cccaggatca gctcatcagg aactcaggat ctggtacctg cctgacatcc	1800
caggacaaaa agccagccat ggccccctgc aatcccagtg acccccatca gttgtggctc	1860
tttgtctagg acccagatca tccccagaga gagccccac aagctcctca ggaaacagga	1920
ttgtgatgt ctgggaacct gatcaccagc ttctctggag gccgtaaaga tggatttcta	1980
aaccacttg gtggcaaggc aggaccttcc taatccttgc aacaacattg ggcccatttt	2040
ctttccttca caccgatgga agagaccatt aggacatata tttagcctag cgttttcctg	2100
ttctagaaat agaggctccc aaagtaggga aggcagctgg gggaggggttc agggcagcaa	2160
tgctgagttc aagaaaagta cttcaggctg ggcacagtgg ctcatgcctg aaatcctagc	2220
actttgggaa gacaatgtgg gagaatggct tgagcccagg agttcaagac cggcctgagc	2280
aacatagtga ggatcccatc tctacgcca ccctcccccc ggcaaaaaaa aaagctgggt	2340
atggtggctt atgcctgtag tcgcagctac tcagaaggct gaggtgggag gattgcttgt	2400
tccccggagg ttgaagctac agtgagcctt gattgtgtca ctgcaactcca gcctgggcaa	2460
caggtaaagac tctgtctcaa aaaaaaaca aaaaaagaaga agaaaagtac ttctacagcc	2520
atgtcctatt ccttgatcat ccaaagcacc tgcagagtcc agtgaaatga tatattctgg	2580
ctgggcacag tggtcacac ctgtaatcct agcacttttg gaggcccaagg caggtggatc	2640
acctgaggtc agaagtttga aaccagcctg gactacatgg tgaaactcca tctctactaa	2700
aagtacaaaa attagctggg catgatggca cgcacctgca gtcccagcta cttgggaggc	2760
tgaggcagga gaatcactcg aaccaggag gcagaggttg cagtgagcca agacagcacc	2820
attgcacccc agcctgagca acaagagcga aactccatct caggaaaaaa aaaaaaaaaa	2880
a	2881

<210> 97

<211> 1930

<212> DNA

<213> NM\_020183.3| Homo sapiens aryl hydrocarbon receptor nuclear translocator-like 2 (ARNTL2), mRNA

<400> 97

gaccaagtgg ctccctgcgat ggcggcgga gaggaggctg cggcgggagg taaagtgttg	60
---	----

agagaggaga accagtgc	tgctcctgtg gtttccagcc	gcgtgagtcc agggacaaga	120
ccaacagcta tggggtcttt	cagctcacac atgacagagt	ttccacgaaa acgcaaagga	180
agtgattcag acccatcccc	gtcaggaatc atgacagaaa	aagtgggtgga aaagctttct	240
cagaatcccc ttacctatct	tctttcaaca aggatagaaa	tatcagcctc cagtggcagc	300
agagtggaag atggtgaaca	ccaagttaaa atgaaggcct	tcagagaagc tcatagccaa	360
actgaaaagc ggaggagaga	taaaatgaat aacctgattg	aagaactgtc tgcaatgatc	420
cctcagtgc accccatggc	gcgtaaaactg gacaaactta	cagttttaag aatggctgtt	480
caacacttga gatctttaaa	aggcttgaca aattcttatg	tgggaagtaa ttatagacca	540
tcatttcttc aggataatga	gctcagacat ttaatcctta	agactgcaga aggcttctta	600
tttgtggttg gatgtgaaag	aggaaaaatt ctcttcgttt	ctaagtcagt ctccaaaata	660
cttaattatg atcaggctag	tttgactgga caaagcttat	ttgacttctt acatccaaaa	720
gatgttgcca aagtaaagga	acaactttct tcttttgata	tttcaccaag agaaaagcta	780
atagatgcca aaactgggtt	gcaagttcac agtaatctcc	acgctggaag gacacgtgtg	840
tattctggct caagacgatc	ttttttctgt cggataaaga	gttgtaaaat ctctgtcaaa	900
gaagagcatg gatgcttacc	caactcaaag aagaaagagc	acagaaaatt ctatactatc	960
cattgcactg gttacttgag	aagctggcct ccaaattattg	ttggaatgga agaagaaagg	1020
aacagtaaga aagacaacag	taattttacc tgccttggtg	ccattggaag attacagcca	1080
tatattgttc cacagaacag	tggagagatt aatgtgaaac	caactgaatt tataaccg	1140
tttgagtgga atggaaaatt	tgtctatgta gatcaaaggg	caacagcgat tttaggatat	1200
ctgcctcagg aacttttg	ggaacttctgt tatgaatatt	ttcatcaaga tgaccacaat	1260
aatttgactg acaagcaca	agcagttcta cagagtaagg	agaaaatact tacagattcc	1320
tacaaattca gagcaaaa	aga tggctctttt gtaacttta	aaagccaatg gtttagtttc	1380
acaaatcctt ggacaaa	aga actggaatat attgtatctg	tcaacacttt agttttg	1440
catagtgagc ctggaga	agc atcattttta ccttgtagct	ctcaatcatc agaagaatcc	1500
tctagacagt cctgtatg	ag tgtacctgga atgtctactg	gaacagtact tgggtgctggt	1560
agtattggaa cagatattgc	aaatgaaatt ctggatttac	agaggttaca gtcttcttca	1620
taccttgatg attcgagtcc	aacaggttta atgaaagata	ctcactactgt aaactgcagg	1680
agtatgtcaa ataaggagt	t gtttccacca agtccttctg	aaatggggga gctagaggct	1740
accaggcaaa accagagtac	tgttgctgtc cacagccatg	agccactcct cagtgatggt	1800
gcacagttgg atttcgatgc	cctatgtgac aatgatgaca	cagccatggc tgcatttatg	1860
aattacttag aagcagaggg	gggcctggga gaccctgggg	acttcagtga catccagtg	1920
accctctagc			1930

<211> 2128

<212> DNA

<213> NM\_014576.2| Homo sapiens apobec-1 complementation factor (ACF),  
transcript variant 1, mRNA

<400> 98  
tttgatatga cgattagagc ataacccgag tgacacgttg aattcgccat aatcaaggaa 60  
accttttccg ggtggggatc tctgaaatta ctgagataac agtgctgtgc caaaaacctg 120  
tggatttttct ctacaaaaat tattgagcaa ccctaattaa cctgattttt tgctgataat 180  
cactctcaat ggaatcaaat cacaaatccg gggatggatt gagcggcact cagaaggaa 240  
cagccctccg cgcactgggtc cagcgcacag gatatagctt ggtccaggaa aatggacaaa 300  
gaaaatatgg tggccctcca cctgggttggg atgctgcacc ccctgaaagg ggctgtgaaa 360  
tttttattgg aaaacttccc cgagaccttt ttgaggatga gcttatacca ttatgtgaaa 420  
aaatcggtaa aattttatgaa atgagaatga tgatggattt taatggcaac aatagaggat 480  
atgcatttgt aacattttca aataaagtgg aagccaagaa tgcaatcaag caacttaata 540  
attatgaaat tagaaatggg cgcctcttag gggtttgtgc cagtgtggac aactgccgat 600  
tatttgttgg gggcatccca aaaacaaaaa agagagaaga aatcttatcg gagatgaaaa 660  
aggttactga aggtgttgtc gatgtcatcg tctacccaag cgctgcagat aaaacaaaaa 720  
accgaggctt tgccttcgtg gagtatgaga gtcacgcagc agctgccatg gcgaggagga 780  
aactgctacc aggaagaatt cagttatggg gacatggtat tgcagtagac tgggcagagc 840  
cagaagtaga agttgatgaa gatacaatgt cttcagtga aatcctatat gtaagaaatc 900  
ttatgctgtc tacctctgaa gagatgattg aaaaggaatt caacaatatc aaaccaggtg 960  
ctgtggagag ggtgaagaaa attcgagact atgcttttgt gcacttcagt aaccgagaag 1020  
atgcagttga ggctatgaaa gctttaaatg gcaagggtgct ggatgggtcc ccattgaag 1080  
tcaccctagc aaaaccagtg gacaaggaca gttatgttag gtatacccga ggcacaggtg 1140  
gaaggggcac catgctgcaa ggagagtata cctactcttt gggccaagtt tatgatccca 1200  
ccacaacctt ccttggagct cctgtcttct atgccccca gacctatgca gcaattccca 1260  
gtcttcattt cccagccacc aaaggacatc tcagcaacag agccattatc cgagcccctt 1320  
ctgttagagg ggctgcggga gtgagaggac tgggcggccg tggctatttg gcatacacag 1380  
gcctgggtcg aggataccag gtcaaaggag acaaaagaga agacaaactc tatgacattt 1440  
tacctgggat ggagctcacc ccaatgaatc ctgtcacatt aaaaccccaa ggaattaaac 1500  
tcgctcccca gatattagaa gagatttgtc agaaaaataa ctggggacag ccagtgtacc 1560  
agctgcactc tgctattgga caagacaaa gacagctatt cttgtacaaa ataactattc 1620  
ctgctctagc cagccagaat cctgcaatcc accctttcac acctccaaag ctgagtgcct 1680  
ttgtggatga agcaaagacg tatgcagccg aatacaccct gcagaccctg ggcatcccca 1740

ctgatggagg	cgatggcacc	atggctactg	ctgctgctgc	tgctactgct	ttcccaggat	1800
atgctgtccc	taatgcaact	gcacccgtgt	ctgcagccca	gctcaagcaa	gcggtaaccc	1860
ttggacaaga	cttagcagca	tatacaacct	atgaggtcta	cccaactttt	gcagtgactg	1920
cccgagggga	tggatatggc	accttctgaa	gatgcttttt	taaatttaag	aataagacac	1980
acaaaactct	attaaaaaaaa	aaaaagaaat	aaacctctaa	ctcgggtcccc	aatgatcata	2040
aataatatgt	ttcctaaaga	aatgcctttc	cagagactgt	atagcttata	ccaattatag	2100
aatcatgaag	taaaaaaaaa	aaaaaaaaa				2128

<210> 99

<211> 5730

<212> DNA

<213> NM\_019008.4| Homo sapiens hypothetical protein FLJ20232 (FLJ20232), mRNA

<400> 99		
cctccgcctc	cactcgccct	cgtgctccct
tcagccccct	cgcagctccg	tgcgcaaggt
		60
cgtgtccccg	aagtgaagg	gccatgttga
tgggtgaccc	ggggagaggt	acccggccag
		120
aggcgagtcc	tgcgagtg	tagcgcgcac
ggcctgcggt	gtgacacca	gccccgtcca
		180
gtcccccatg	gccccgtgga	gccgagaggc
gggtgctgagt	ctctatcggg	ctctgttgcg
		240
ccagggccga	cagcttcgct	acactgatcg
agacttctac	tttgccctcca	tccgccgtga
		300
attccgaaaa	aatcagaagc	tagaggacgc
tgaggccccg	gagaggcagc	tggagaaggg
		360
cctgggtcttt	ctcaacggca	aattggggag
gatcatttag	gatcctccaa	gggaaagagg
		420
acaaaggtgc	cttctgtaga	cactcctgct
ctcttccatc	cccatcttac	agatgtatta
		480
agaagcctca	gatgagcaat	ggcaggcgct
ggtgagcgca	aaggcaagaa	ggatgacaat
		540
ggcattggca	cggccattga	ctttgtgctc
tccaatgccc	ggctggtgct	gggggtgggt
		600
ggagcgcca	tgctgggcat	cgccacgctg
gcagttaagc	ggatgtacga	tcgggcatc
		660
agtgccccta	ccagccccac	ccgcctgagc
cattcgggga	aaaggagctg	ggaagaaccc
		720
aactggatgg	gctccccacg	actgctgaac
agggacatga	agacgggcct	gagccggtcc
		780
ttgcagaccc	ttcccacaga	ctcctccacc
ttcgacacag	atacattctg	cccgccccgg
		840
cccaagccag	tggccaggaa	gggcccaggta
gacttgaaga	agtcacgact	ccgcatgtcc
		900
ctgcaggaga	aacttcttac	ttactaccgg
aaccgggcag	ccatccctgc	tggagagcag
		960
gctcgggcca	agcaagctgc	tgtggacata
tgtgccgagc	tccggagctt	cctgcggggc
		1020
aagttgcctg	acatgccgct	tcgggacatg
tacttgagtg	gcagcctcta	cgatgacctg
		1080
cagtggtgta	cagctgacca	catccaactc
attgtgcccc	ttgtgctgga	gcagaacctg
		1140
tggtcatgta	ttcctgggtga	agacaccatc
atgaatgtcc	ctggcttctt	cctggtgcgt
		1200

cgtgagaatc	cagagtactt	tcctcgtggg	agcagttact	gggaccgctg	tgtagtaggg	1260
ggctacctct	ctccaaagac	agtcgcagat	acattttgaga	aggtagtggc	tggctccatc	1320
aattggccag	ccataggggtc	cctcttggac	tatgtgatcc	gcccggcccc	acccccagaa	1380
gcccctcacac	tggaggtgca	gtatgagcgt	gacaaacatc	tcttcattga	cttcctgcca	1440
tcagtgaccc	tcggtgacac	agtcttgggtg	gccaaccac	accggctagc	ccagtatgac	1500
aacctgtggc	ggctgagcct	gcgtcccgcg	gagacggcac	gcctgcgggc	tctggaccag	1560
gctgactcgg	gctgccgac	tctgtgcctc	aagatcctca	aggccatatg	caagtccacc	1620
ccggctctgg	gccacctcac	tgccagccag	ctaaccaatg	tcatectcca	cttggcccag	1680
gaggaggctg	actggtctcc	ggatatgctg	gccgaccgtt	tcctgcaggc	cttgagggga	1740
cttatcagct	acttagaggc	tggagtcctg	cccagtgccc	taaaccccaa	gggtgaactta	1800
tttgagagc	tcacccctga	agaaatagac	gaattaggat	acactctgta	ttgctcattg	1860
tctgagccag	agggtgctgct	gcagacgtag	ggcaggtgaa	ggccaaagcg	gggtgttggtg	1920
gtcaggccct	ggattctccg	ttagatacac	ttggctacct	agttggtgcc	tcacagggtt	1980
cctgctgcct	gggtgtcttg	tgatcatcac	cctggtcact	tcatgctgat	tagaatgaca	2040
tctctttcgt	ctcctatfff	gttaccacaac	tcttcctatt	tttgttacca	atcactgtgc	2100
tctctgccgc	cccctggctc	caggctaatt	tttctggaat	gaattgagaa	gggtggcgtgc	2160
tggcctgagc	tgatggacca	cttgggtgtt	tgcgttttgg	cccatgtttg	ctgcctctat	2220
ctggtctgcc	ttgcccgttt	gcctgttcc	attcagtgtc	ttttctatff	tttcctctct	2280
cgttcatgcc	ttctgttttg	ctcttgtccc	tggagcatat	ctgcctaatt	aagatgttgc	2340
cttttagttg	aatgccactg	aagagctgtg	atagcatgtt	tcaaagctga	actctacaga	2400
gcgagtgtg	agacagtatt	tagggtttct	gggagtgagg	ctggtagaag	agttggcctt	2460
tgaccacggt	tcctggagta	gaagtcacac	ctcccccaa	cctcctgacc	cattcataaa	2520
tgctgagaat	gtctctcatg	ggaacactgt	taatgacca	cacaggataa	gctgaatgca	2580
aagttatttg	caggttgaat	ttcttgggtg	ctattagcag	aagtgcagag	tagggaacca	2640
gagctgggta	agggcctagt	gaagggtttg	tgtgcccagt	gtctgctcgt	catctgtggc	2700
tgcaggggtc	agacagacaa	ggatggggac	tgccagggca	ccacttcac	atgaatgctg	2760
gttttcacac	cttttcctta	ttttattgcc	aatcaggaca	aggccttgaa	ggaacgcagc	2820
cttagacatc	agggtaggat	gatggaggta	gacagtcgac	tgaatgtcag	ctggaaaatc	2880
cagtcactag	ttgggggtttg	gtggccatgt	tttctacca	gacaggccct	gcttttctag	2940
gatgtggcct	tagagcaaga	acagacccaa	cagccagccc	ttcatcctcc	agcgtctgcc	3000
ataggaatgt	gagaggggtg	tttgctgagc	gctccgggca	cggccagagg	gcaagtgagc	3060
atgcacggac	ctcttcccc	tgtcctgttt	ctcaccagc	acctggggag	atcgggtgcta	3120
ccaaggaaga	gagcacacag	ataagacaga	ggggaggagg	tgggcatttc	ctacattcct	3180
ccttgtttgc	cgctgctgag	attgcagtat	ttattgcaat	gtaaatgtat	cctgaagggtg	3240

gggaggaatg	tttaatctac	catgtccgtg	tgtcatcttg	gtttgtgttt	ttccctgttt	3300
gtagcaagac	tctgatgata	attctgtttc	tcatctgccc	attcagtatt	ttgttttcct	3360
tccgtcaagt	tgtcttattt	tttcaatgac	tacctctcca	tcattgaggt	tctggtgaag	3420
ctctctgcag	ctgtctcatt	ccttcccaac	gatagtaaca	ggaaatgact	ctttagcatc	3480
gatacctcaa	catcaattta	gggtagagat	tcctgcccct	cttttgtcac	agattaggaa	3540
attgagaact	agggttaacc	ttgactatat	ttagaggtct	ttttgcctct	tttcccctta	3600
acaaggattt	cttatggtgg	tttcagtttc	atttgcataa	aggtattgag	agggaacaaa	3660
aaacataaag	ctgagaatct	tgagagagct	catctaccct	gtctgttggg	cagactcaaa	3720
tgagagttaa	aaaaaaaaaa	aaaaatctgt	atgcctgagt	accatcctgg	atgaatctag	3780
aaggtatggg	gtagagcttg	acagggttcc	tgtgtaccca	ctgggtatcc	gttagaggta	3840
agggagagga	gaggattgat	agagtgttgc	aaaagtatag	attattcatt	gagataaagg	3900
atttggtttc	cctgccatga	gtattaaaaa	aatttaagtt	ttcccaagct	tgcatctctg	3960
accaaatttc	acataaaaaca	ttggaaggag	gctgggtgcg	gtggctcatg	cttgtaatcc	4020
cagcactggg	aagctaaggc	gggtggatca	cttgagggtca	ggagtccgag	accagcctgg	4080
ccaacatggt	gaaaccccg	ctccacgaaa	aagataaaaa	taagctgggc	gtggtggcag	4140
gcgcctataa	tcccagctac	tccggagggt	gaggcaggag	aataacttaa	acccgggagg	4200
cggaggttac	agtgagctga	gatcgtgcc	ctgcactcca	gcctgggtga	cagagtgaga	4260
ccctattttca	aaaaaataaa	aattggaaga	agagcttaaa	aaagataaga	ttttaagag	4320
tcccaagtta	tttaagttga	gtgtaattgt	catttaagga	aggcaaata	gtttatcatc	4380
cttcttaaaag	agcatctctt	ttaactgttg	gacaaaacca	taactttgtc	attttacaag	4440
gaagaacctc	ttaagaagtc	ctcagaacca	gaagcaatgt	gaactctcag	cgctggctct	4500
gggtgggtttg	ctgaccatga	ctgggcaagc	cgttcttttt	gctgccatct	tcctcatcat	4560
aaagtgtgga	acataggcaa	ttgctttgag	attcttggat	agaagaggac	aacattctgc	4620
acctgcccc	ttttttaaat	ctttggggaa	agatgagtaa	ctttccccac	tactctgcct	4680
tcctgttcag	taactcttac	ttttgcctga	agtaacagca	tcttctactt	ctccatctag	4740
agatttttgt	gtgtgtgcca	tcaagggttag	caaactttat	acgtagccta	acacttaaaa	4800
aatgcactca	ttatcttaaa	cctaataaat	tccagagttt	attttggttc	tcctctgttg	4860
cccttcctaa	aaaatgagct	gaagatgaca	gtatttttct	ttacatgctt	ggttatgact	4920
tttaaagttt	tatttaaata	aatgttgaag	ctcaagttta	aagaagcggt	gcagaggccc	4980
acggtctcct	gggtcccggc	cacctgtcca	tattccacat	ttgctgactg	tgctccctgc	5040
actccactca	agttgagagt	tcaaataagtc	ttgaagggga	atcagcttca	ggatggaagg	5100
accaggaga	ggccccgagg	tgggaggggt	ctgtaaatac	agactactgc	gagtgtccag	5160
agctctctgc	catgatactt	ccttgggact	gacttggctg	agaacgtgtt	ctgtcagagg	5220



atttgttaga	actctgccct	tttgtctgaa	actcaaggcc	aaggagaatg	ataggagact	5280
taggacagag	ctgacccttg	caccaggctg	ggaggctgca	gcccttttag	atgccactta	5340
ctgtaagtgg	ccagaatacc	agagaggtgg	gttccatggt	caaatgcaca	gtaggtgttt	5400
acctttacat	ttggatcacc	ttgtagtctt	taaattcttg	gtccctgagg	ccaagtccac	5460
aacttgccct	ctagtcactt	gcctgcccgc	agtggtggtg	gatgtgttag	ctggtagatt	5520
tggaatcagt	caccagtctt	tctgtactgt	cttggttagc	tctatataag	taggggcagc	5580
ttagccctga	ggcccagaga	cctgctgtcc	tttttctcct	tgagggagga	aataaaaactg	5640
cggaatacaa	tgtccttcca	tagcatggga	agaagaaaat	aaacatctcc	ttccaacaa	5700
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa				5730

<210> 100

<211> 2545

<212> DNA

<213> NM\_030882.2| Homo sapiens apolipoprotein L, 2 (APOL2), transcript variant alpha, mRNA

<400> 100						
gtgctgggga	gcagcgtgtt	tactgtgctt	ggcatgagc	tgctgggaag	ttgtgacttt	60
cactttccct	ttcgaattcc	agggtatatc	tgggaggccg	gaggacgtgt	ctggttatta	120
cacagatgca	cagctggacg	tgggatccac	acagctcaga	acagttggat	cttgctcagt	180
ctctgtcaga	ggaagatccc	ttggacaaga	ggaccctgcc	ttggtgtgag	agtgagggaa	240
gaggaagctg	gaacgagggg	taaggaaaac	cttccagtct	ggacagtgac	tggagagctc	300
caaggaaagc	ccctcggtaa	cccagccgct	ggcaccatga	acccagagag	cagtatcttt	360
attgaggatt	accttaagta	tttccaggac	caagtgagca	gagagaatct	gctacaactg	420
ctgactgatg	atgaagcctg	gaatggattc	gtggctgctg	ctgaactgcc	cagggatgag	480
gcagatgagc	tccgtaaagc	tctgaacaag	cttgcaagtc	acatgggtcat	gaaggacaaa	540
aaccgccacg	ataaagacca	gcagcacagg	cagtggtttt	tgaaagagtt	tcctcggttg	600
aaaaggggagc	ttgaggatca	cataaggaag	ctccgtgccc	ttgcagagga	ggttgagcag	660
gtccacagag	gcaccaccat	tgccaatgtg	gtgtccaact	ctgttggcac	tacctctggc	720
atcctgaccc	tcctcggcct	gggtctggca	cccttcacag	aaggaatcag	ttttgtgctc	780
ttggacactg	gcatgggtct	gggagcagca	gctgctgtgg	ctgggattac	ctgcagtgtg	840
gtagaactag	taaacaaatt	gcgggcacga	gcccagccc	gcaacttgga	ccaaagcggc	900
accaatgtag	caaagggtgat	gaaggagttt	gtgggtggga	acacacccaa	tgttcttacc	960
ttagttgaca	attggtacca	agtcacacaa	gggattggga	ggaacatccg	tgccatcaga	1020
cgagccagag	ccaaccctca	gttaggagcg	tatgccccac	ccccgcatat	cattgggcga	1080

```

atctcagctg aaggcgggtga acagggttgag aggggttggtg aaggccccgc ccaggcaatg 1140
agcagaggaa ccatgatcgt ggggtgcagcc actggaggca tcttgcttct gctggatgtg 1200
gtcagccttg catatgagtc aaagcacttg cttgaggggg caaagtcaga gtcagctgag 1260
gagctgaaga agcgggctca ggagctggag gggaagctca actttctcac caagatccat 1320
gagatgctgc agccaggcca agaccaatga cccagagca gtgcagccac cagggcagaa 1380
atgccgggca caggccagga caaaatgcag actttttttt tttttttttt ttttttttga 1440
gatggagtct cgctctatcg cccaggatgg agtgcagtgg ctcaatctcg gctcactgca 1500
aactccgcct cccgggttca caccattctc cggcctcagt ctcccagta gctgggacta 1560
caggcacctg ccaccacgcc cggctaattt ttttgtattt tctactggaga cggggtttca 1620
ctgtgttagc cacgatggtc tccatctcct gacctcgtga tctgcccacc tcggcctccc 1680
aaagtgctgg gattacaggc gtgagccacc gcgcctggcc aaaatgcaga cattttatta 1740
gggggataag gagggcaagg taaagcttat ggaactgagt gttagtact ttggcatttg 1800
tgtagctgag cacagcaagg gaggggttaa tgcagatggc aagtgcacca aggagaaggc 1860
aggaacactg gagcctgcaa taaggaggga gagaggactg gagagtgtgg ggaatgggaa 1920
gaagtagttt actttggact aaagaatata ttgggcgaag aatagagggg gagcttgtag 1980
gaaccagcaa tgagaaggcc aggaaaagaa agagctgaaa atggagaaaa ccagagttag 2040
aactgttggg tacaggagaa gaaacagcag ctccactacc gacccccccc caggtttgat 2100
gtccttccaa gaataaagtc tticcctggg gatggtctct cgctctgtct ttccagcatc 2160
cactctccct tgccttctg ggggtgtatc acagtcagcc agtggcttct tcatgatggg 2220
ggttgggggtg gttgtcatgt gacgggtccc ctccaggtta ctaaagggtg catgtccctt 2280
gcttgaacct tgagaggcag gtggtaggcc atggccacaa tcccagctg aggagcaggt 2340
gtccctgaga acccaaactt cccagagagt atctgagaac caaccaatga aaacagtccc 2400
atcgctctta gccggtagt aaacagtcag aagattagca tgaaagcagt ttagcattgg 2460
gaggaagcac agatctctag agctgtcctg tcgctgcca ggattgacct gtgtgtaagt 2520
ccaataaac tcacctactc accaa 2545

```

<210> 101

<211> 1429

<212> DNA

<213> NM\_016612.1| Homo sapiens mitochondrial solute carrier protein (MSCP), mRNA

<400> 101

```

ccccctcccc tcctgcagcc tcctgcgccc cgccgagctg gcggatggag ctgcgcagcg 60
ggagcgtggg cagccaggcg gtggcgcgga ggatggatgg ggacagccga gatggcggcg 120

```

gcggcaagga	cgccaccggg	tcggaggact	acgagaacct	gccgactagc	gcctccgtgt	180
ccaccacat	gacagcagga	gcgatggccg	ggatcctgga	gcactcggtc	atgtacccgg	240
tggactcgg	gaagacacga	atgcagagtt	tgagtccaga	tcccaaagcc	cagtacacaa	300
gtatctacg	agccctcaag	aaaatcatgc	ggaccgaagg	cttctggagg	cccttgcgag	360
gcgtcaacgt	catgatcatg	ggtgcagggc	cggcccatgc	catgtatttt	gcctgctatg	420
aaaacatgaa	aaggacttta	aatgacgttt	tccaccacca	aggaaacagc	cacctagcca	480
acgggatagc	tgggagtatg	gccaccctgc	tccacgatgc	ggtaatgaat	ccagcagaag	540
tgggtgaagca	gcgcttgag	atgtacaact	cgcagcaccg	gtcagcaatc	agctgcatcc	600
ggacggtgtg	gaggaccgag	gggttggggg	ccttctaccg	gagctacacc	acgcagctga	660
ccatgaacat	ccccctccag	tccatccact	tcatcaccta	tgagttcctg	caggagcagg	720
tcaacccccca	ccggacctac	aaccgcaggt	cccacatcat	ctcaggcggg	ctggccgggg	780
ccctcgccgc	ggccgccacg	acccccctgg	acgtctgtaa	gacccttctg	aacactcagg	840
agaacgtggc	cctctcgctg	gccaacatca	gcggccggct	gtcgggtatg	gccaatgcct	900
tccggacggt	gtaccagctc	aacggcctgc	cggctacttc	aaaggcatcc	aggcgcgtgt	960
catctaccag	atgccctcca	ccgccatttc	ttggtctgtc	tatgagttct	tcaagtactt	1020
tctcaccaag	cgccagctgg	aaaatcgagc	tccatactaa	aggaaaggat	catagaatct	1080
tttcttaaaag	tcattctctg	cctgcatcca	gccccctgcc	ctctcctcac	acgtagatca	1140
tttttttttt	tgcaggggtgc	tgcctatggg	ccctctgctc	cccaatgcct	tagagagagg	1200
aggggacggc	acggccgctc	accggaaggc	tgtgtgcggg	gacatccgag	gtggtgggtg	1260
acaggaagga	cttgggaagg	ggagcgagaa	attgcttttt	ctcttcctcc	ctgggcagaa	1320
tgtagctttt	ctgcttcact	gtggcagcct	cctccctgga	tccttagatc	ccagaggagg	1380
gaagaaaatt	tgcagtgact	gaaaacagta	aaaaaaaaaa	aaaaaaaaaa		1429

<210> 102

<211> 2368

<212> DNA

<213> NM\_017903.2| Homo sapiens hypothetical protein FLJ20618 (FLJ20618), mRNA

<400> 102	
ccacgcgtcc	gaaaatgttt gaacaattgg atttcaaaca ttttcgtttt gtggagtgggt 60
gctcaccaag	tggtagagcc ctaagcaagt gaacacaaac acatttaagt gtattttgtc 120
tgattagatg	ttagccagtt atgctatttc attcaaagt ctgaaaaaat caattgacta 180
ttcccttttc	ctaaagggca gagacagata atctcacttc cagagaaatg acttggagaa 240
aaaaaagtgt	tggctttttt gctcttttgt aattaaatcc ggatgtacct caaaagactt 300
aagactgtgg	tgataagatg ctttcctcag cagaaaggag ggaaaaaaaa acaactggaa 360

ctcaaagctt gaaattctgt ggcaaaacat gagatgtcca ggattggagg ttgaaaagat	420
ttcactacag tgttctgcaa tagttggagc agataacttt cagtgtagcc acagccatgg	480
actccagatt tccagatttt caagacctgg acctggaacc cgaaagagct tgtcacgatg	540
cggcaggaac actggaggta gatttttttt tatttttgaa ttttgggact gttgaccttg	600
ctgtgagaaa agagacaacg actgagcaag cactaccacc agcactgtta ctgggaatta	660
gaagacctga gtttctgtcc agaccctcag tgcaaaactga ggatgctcca tccaaagtga	720
attatgtcct gtgcctcctg attgctgagt gttcacctgg accttctgac taccttccct	780
gtgctattcc atcagcctac agacctggta cctggatttt tgcccagat gattcctacc	840
accttactac tgacgaagac acccattcca gtggaccact gtgaccagg aggcattcag	900
ccatcatgat gtggccttta cctccactcc tgtcttgttc taccagatt cagcacagcc	960
ctttatagtg aagtcagagt cctcaagcca aatagctaaa gctgttttat cacaacaaag	1020
gcctagtttg ttccatgagt gtgcatttca tttcttcagt taaagccttc agagacacac	1080
aataaatttg gaccagggga ttttttagtt attaattgctc tctgaagaaa ggcaacatct	1140
ttttgagagc agcattggac cacacccac aatctcaaat gattgaaatt catgaacatc	1200
taggatcccg tgaaggtcac tggaccctgt ttttctact tcaaactctg tagtagccta	1260
ctgaatgaga aaacatatte tgaccattg ggatcaaate aaaggcacag tgaactcctc	1320
atagcatctt ctttggaatt actcaggaac cagaactttt tacacaaatg taagaaattc	1380
taccaaggag tccccttacc taacagcatc tcacaaggct gcaccagatt ccagaaaagg	1440
cttctcttga tacatcaagc attttgtgac cgacttattc ttagatcatt ggttttccaa	1500
aggctttgtg gccatgaagc cctttgagtg aaaactgtgc agaagcccag agtaaaagtg	1560
aagctgctct ggatgaagta gtgaagcaag agtaggggcc tgaatcctgc tacaactatc	1620
ttcctttacc accgtggtga cacctaaggg gacttcctta caacacctg aactcttccg	1680
aacacagttt gaaaaccact gccccagaca gcaatatgtt tgacctgaat ggcatccaa	1740
tcttttctgt acctccactc agcacagttc atgttcagta gatgctgaac attcttagaa	1800
atactgtgtg tgaacttaga aaagtgcaag aagacaggca tgtctttgac cccaggaatg	1860
atcatttgct gaagatggtg tcaagtgaac ctagattaac agccctccac tccagatgga	1920
tatccagtga ttcctagaat gggatatagc cagagaacaa ttctatgcac cctacactga	1980
cagactccct taagcaacac cagatgctct actggtactt gaagtacatg actttgaagt	2040
cttgaccctc catgaatacc tgaattatca gcaagcgggt tttgaagctg gtgcctcatt	2100
gaggccatat tagagcaact tgtacatttg acctcttggt atcagccatg gtactctact	2160
tcgtgtgcaa gagataacta tgaaagccaa attcaaatac tggcaacatt tcctaaaggg	2220
gctcaatatac tatcattcgt ctcttttcc aaactacaca tcaactgtatg actcaaccag	2280
tagcagttat attgcccctt ggtttttatt cagttaact actgtttcca agataaatga	2340

gctaataagc tttaaaaaaaa aaaaaaaaa

2368

<210> 103

<211> 2577

<212> DNA

<213> nm\_003011.1 SET translocation (myeloid leukaemia-associated) Homo sapiens

<400> 103  
cacatgtcgg cgcaggcggc caaagtcagt aaaaaggagc tcaactccaa ccacgacggg 60  
gccgacgaga cctcagaaaa agaacagcaa gaagcgattg aacacattga tgaagtacaa 120  
aatgaaatag acagacttaa tgaacaagcc agtgaggaga ttttgaaagt agaacagaaa 180  
tataacaaac tccgccaacc attttttcag aagaggtcag aattgatcgc caaaatccca 240  
aatttttggg taacaacatt tgtcaacat ccacaagtgt ctgcactgct tggggaggaa 300  
gatgaagagg cactgcatta tttgaccaga gttgaagtga cagaatttga agatattaaa 360  
tcaggttaca gaatagattt ttattttgat gaaaatcctt actttgaaaa taaagttctc 420  
tccaaagaat ttcattctgaa tgagagtggg gatccatctt cgaagtccac cgaaatcaaa 480  
tggaatctg gaaaggattt gacgaaacgt tcgagtcaaa cgcagaataa agccagcagg 540  
aagaggcagc atgaggaacc agagagcttc ttacctggg ttactgacca ttctgatgca 600  
ggtgctgatg agttaggaga ggtcatcaaa gatgatattt ggccaaaccc attacagtac 660  
tacttggttc ccgatatgga tgatgaagaa ggagaaggag aagaagatga tgatgatgat 720  
gaagaggagg aaggattaga agatattgac gaagaagggg atgaggatga aggtgaagaa 780  
gatgaagatg atgatgaagg ggaggaagga gaggaggatg aaggagaaga tgactaaata 840  
gaacactgat ggattccaac cttccttttt ttaaattttt tccagtccct gggagcaagt 900  
tgcagtcttt tttttttttt tttttttttt ccctcttggt ctcagtcgcc ctgttcttga 960  
ggtctctttt ctctactcca tggttctcaa tttatttggg gggaaatacc ttgagcagaa 1020  
tacaatggga aaagagtctc taccctttt tggtcgaagt tcatttttat cccttcctgt 1080  
ctgaacaaaa actgtatgga atcaacacca ccgagctctg tgggaaaaaa gaaaaacctg 1140  
ctccctttgc tctgctggaa gctggagggt gctaggcccc tgtgtagtag tgtatagaat 1200  
tctagctttt ttctctctt ctctgtatat tgggctcaga gagtacactg tgtctctatg 1260  
tgaatatgga cagtttagcat ttaccaacat gtatctgtct actttctctt gtttaaaaaa 1320  
agaaaaaaa acttaaaaaa atgggggttat agaaggtcag caaaggggtg gggtttgaga 1380  
tgtttgggtg ggtagtggg cattttgaca acatggcttc tcctttggca tgtttaattg 1440  
tgatatttga cagacatcct tgcagtttaa gatgacactt ttaaaataaa ttctctccta 1500  
atgatgactt gagccctgcc actcaatggg agaatcagca gaacctgtag gatcttattt 1560

ggaattgaca	ttctctattg	taattttggt	cctgtttatt	tttgggtttc	tttttgtttc	1620
actggaaagg	aaagatgatg	ctcagtttta	aacgttaaaa	gtgtacaagt	tgctttgtta	1680
caataaaact	aaatgtgtac	acaaaggatt	tgatgctttt	ctctcagcat	aggtatgctt	1740
actatgacct	tccaagtttg	acttgtataa	catcactgtc	aaactttgtc	accctaactt	1800
cgtatTTTTT	gatacgcact	tttgcaggat	gacctcaggg	ctatgtggat	tgagtaatgg	1860
gatttgaatc	aatgtattaa	tatctccata	gctgggaaac	gtgggttcaa	tttgccattg	1920
gtttctgaaa	agtattcaca	tcattttgga	taccagatag	ctcaatactc	tctgagtaca	1980
ttgtgccctt	gattttttatc	tccaagtggc	agtttttaaa	attggccttt	tacctggata	2040
taaattaatt	gtgcctgcc	ccaccatcca	acagacctgg	tgctctaata	ccaagttata	2100
cacgggacag	ttgctggcat	gtcttcattg	gctctctaaa	atgtggccaa	gaagataggc	2160
tctcagtaag	aagtctgatg	gtgagcagta	actgtccctg	ctttctggta	taaagctctc	2220
aaatgtgacc	atgtgaatct	gggtgggata	atggactcag	ctctgtctgc	tcaatgccat	2280
tgtgcagaga	agcaccctaa	tgcataagct	ttttaatgct	gtaaaaatata	gtcgctgaaa	2340
ttaaatgcc	ctttttcaga	ggtgaattaa	tggacagtct	ggtgaacttc	aaaagctttt	2400
tgatgtataa	aacttgataa	atggaaactat	tccatcaata	ggcaaaaagt	taacaaccta	2460
tctagatgga	tagtatgtaa	tttctgcaca	ggtctctgtt	tagtaaatac	atcactgtat	2520
accgatcagg	aatcttgctc	caataaaagga	acataaagat	ttaaaaaaaa	aaaaaaa	2577

<210> 104

<211> 7577

<212> DNA

<213> XM\_030577.9| PREDICTED: Homo sapiens ATPase, Class II, type 9A (ATP9A), mRNA

<400> 104						
atgacggaca	acatcccgtc	gcagccggtg	cgccagaaga	agcggatgga	cagcaggccc	60
cgcgccgggt	gctgcgagtg	gctgagatgc	tgcggtggag	gggaggccag	gccccgcact	120
gtctggctgg	ggcaccccga	gaagagagac	cagaggtatc	ctcggaatgt	catcaacaat	180
cagaagtaca	atttcttcac	ctttcttcct	ggggtgctgt	tcaaccagtt	caaatacttt	240
ttcaacctct	atttcttact	tcttgccctg	tctcagtttg	ttcccgaat	gagacttggt	300
gcactctata	cctactgggt	tcccctgggc	ttcgtgctgg	ccgtcactgt	catccgtgag	360
gcggtggagg	agatccgatg	ctacgtgcgg	gacaaggaag	tcaactccca	ggtctacagc	420
cggctcacag	cacgaggcac	agtgaagggt	aagagttcta	acatccaagt	tggagacctt	480
atcatcggtg	aaaagaacca	gcgggtccct	gccgacatga	tcttcttgag	gacatcagaa	540
aaaaacgggt	catgcttctt	gcggacggat	cagctggatg	gggagacgga	ctggaagctg	600

cggtttcccg	tggcctgcac	gcagaggctc	cccacggccg	ccgaccttct	tcagattcga	660
tcgtatgtgt	acgcagaaga	gccaaatatt	gacattcaca	acttcgtggg	aactttttacc	720
cgagaagaca	gcgaccccc	gatcagcgag	agcctgagca	tagagaacac	gctgtgggct	780
ggcactgtgg	tcgcatcagg	tactgtttgt	gggtgtgttc	tttactctgg	cagagaactc	840
cggagtgtca	tgaatacctc	aaatccccga	agtaagatcg	gcctgttcga	cttggaagtg	900
aactgcctca	ccaagatcct	ctttggtgcc	ctgggtgggtg	tctcgctggg	catgggtgcc	960
cttcagcact	ttgcaggccg	ttggtacctg	cagatcatcc	gcttcctcct	cttgttttcc	1020
aacatcatcc	ccattagttt	gcgtgtgaac	ctggacatgg	gcaagatcgt	gtacagctgg	1080
gtgattcgaa	gggactcgaa	aatccccggg	accgtgggtc	gctccagcac	gattcctgag	1140
cagctgggca	ggatttcgta	cttactcaca	gacaagacag	gcactcttac	ccagaacgag	1200
atgattttca	aacgggtcca	tctcggaaca	gtagcctacg	gcctcgactc	aatggacgaa	1260
gtacaaagcc	acattttcag	catttacacc	cagcaatccc	aggaccacc	ggctcagaag	1320
ggcccaacgc	tcaccactaa	ggtccggcgg	accatgagca	gccgcgtgca	cgaagccgtg	1380
aaggccatcg	cgctctgcca	caacgtgact	cccgtgtatg	agtccaacgg	tgtgactgat	1440
caggctgagg	ccgagaagca	gtacgaagac	tcctgcccg	tataccaggc	atccagcccc	1500
gatgagggtg	ccctggtaca	gtggacggaa	agtgtgggct	taaccctggg	gggccgagac	1560
cagtcttcca	tgcagctgag	gacccctggc	gaccagatcc	tgaacttcac	catcctacag	1620
atcttccctt	tcacctatga	aagcaaacgt	atgggcatca	tcgtgcggga	tgaatcaact	1680
ggagaaatta	cgttttacat	gaaggagca	gatgtgggtc	tggctggcat	tgtgcagtac	1740
aatgactggg	tggaggaaga	gtgtggcaac	atggcccag	aagggtgctg	ggtgctcgtg	1800
gtggcaaaga	agtctcttgc	agaggagcag	tatcaggact	ttgaagccc	ctacgtccag	1860
gccaagctga	gtgtgcacga	ccgctccctc	aaagtggcca	cggtgatcga	gagcctggag	1920
atggagatgg	aactgctgtg	cctgacgggc	gtggaggacc	agctgcaggc	agatgtgctg	1980
cccacgctgg	agaccctgag	gaatgctggc	atcaaggttt	ggatgctgac	aggggacaag	2040
ctggagacag	ctacgtgcac	agcgaagaat	gcacatctgg	tgaccagaaa	ccaagacatc	2100
cacgtttttc	ggctggtgac	caaccgcggg	gaggctcacc	tcgagctgaa	cgccttccgc	2160
aggaagcatg	attgtgccct	ggtcatctcg	ggagactccc	tggaggtttg	cctcaagtac	2220
tatgagtacg	agttcatgga	gctggcctgc	cagtgcccg	ccgtagtctg	ctgccgatgt	2280
gccccaccc	agaaggcca	gatcgtgcgc	ctgcttcagg	agcgcacggg	caagctcacc	2340
tgtgcagtag	gggacggagg	caatgacgtc	agcatgattc	aggaatctga	ctgcggcgtg	2400
ggagtggaa	gaaaggaagg	aaaacaggct	tcgttggtg	cagacttctc	catcactcaa	2460
tttaagcatc	ttggccgggt	gcttatgggt	catggccgga	acagctacaa	gcggtcagcc	2520
gccctcagcc	agttcgtgat	tcacaggagc	ctctgtatca	gcaccatgca	ggctgtcttt	2580
tcctccgtgt	tttactttgc	ctccgtccct	ctctatcaag	gattcctcat	cattgggtac	2640

tccacaat	ttt acaccatg	ttt tccctgtg	ttt tctctgg	tcc tggacaa	aga tgtcaa	atcg	2700
gaagttg	cca tgctgt	atcc tgagct	tac aaggat	cttc tcaagg	gacg gccgtt	gtcc	2760
tacaag	acat tctta	atag ggtttt	gatt agcat	ctatc aagg	gac catcat	gtac	2820
ggggcg	ctgc tgctgt	tttga gtcgg	agttc gtgc	acatcg tggc	atctc cttcac	ctg	2880
ctgatc	ctca ccgag	ctgct cat	gggtgg	cgc ctgac	catcc agac	ctgg	2940
acagtg	gcgg agctg	ctcag cctgg	cctgc tacat	cgct ccct	gggtgt	cttac	3000
ttcatc	gatg tgtac	ttcat cgcc	accttg tcatt	cttgtt ggaa	agtctc	cgatc	3060
ctggtc	agct gcct	ccccct ctatg	tcctc aagt	acctgc gaag	acgggt	ctctc	3120
agctact	caa agctc	acatc atagg	ccgtg cgttc	gctgg agggg	ccct ggtc	ttgg	3180
cttcctg	at ggac	agagct caagt	ttccat ttat	attaac cgcc	acctgt	ggattt	3240
gtaattg	cta acac	atgcag tttta	atggg aagt	ggctct	gcgc	ctaaac	3300
cgctgc	atca acggg	agggga gggc	ctgaa agag	acccat ctgg	gcctgt	ctga	3360
cgttctt	cat gttt	aggtga atat	gaatat gttaa	agctg gtgg	ctcagc	tggg	3420
atatggg	tca ctgtg	cgcagc ttcct	tatga cttga	atttt gttgt	cacat gata	aaaagt	3480
tctgtg	tagc tga	aggttg aga	aggcttg tgtg	tgtgt	tgtg	tgtg	3540
tgtgtg	tttt taa	agagtc taat	gtgtat gtact	ctttta tgtc	tttctt	gctctt	3600
agaggtg	tca gaaaa	ataga aagct	cttgg tgcg	gtttg ggag	gaaaag	acagt	3660
ttggta	aaaaa gttat	ccaca caata	atctc cattc	ggaaa tgct	cagtat	cgtctc	3720
cagccct	gct tatcc	aggtt aactg	ggatt cctgg	gatcg taacc	agtaa atg	agagg	3780
agggag	agag agtgt	ctctaa gtcca	atctg ttat	ccttga tctg	attcag	catcca	3840
gtgtg	agtta actt	cacctg ccac	ctgta aaag	atttc agag	gttga	tccc	3900
ttggg	acctg gta	acaatca caa	agccagt ggct	gtttga	gaagg	acctc	3960
agcag	agttg tttt	tagcagg aaac	gtgcca ctga	atggcc	cctaa	atgtg	4020
gataa	gagac tca	actaatt cttt	taggcaa catg	gcagat	gtgact	caga	4080
ccaaag	cggga aaggt	caggg ggct	gggact cttc	tcttcc	ataga	agcct	4140
aggagg	cata atg	gaagatg accc	cacaaa ggc	agggca	tcttt	cggaa	4200
ggcag	ctttc aga	acaagga accc	ctggtg ggag	gacgcc	caag	ctacag	4260
tggg	atctgt tcc	actgccg gcag	atttca aggg	gaactt	gctga	aaaggc	4320
gaag	atttct ccc	ctccag gatg	gactac atg	ccggcat	gtttc	ttata	4380
tgctt	gtttc agag	gaaggg agtt	tgcagt cgcg	ggacgt	ggtag	agcaa	4440
ggttt	tcaag ttg	cttctg caga	agccac atat	gcatgc	cata	aggggt	4500
atctt	taaga gcc	aagtgtg gttg	agatct tgg	atttgcg	tttact	tctt	4560
tatcctt	caa acc	ctctgcc tggc	gctact tctg	tgtgt	tctg	atgt	4620



ctggtttctg	atgcctacta	actcctgctc	ttggagagct	ggagacacga	ggatcagata	4680
gtcccttgcc	tttggagcac	tcttgataag	cttttgtatt	ttgtgttgtc	cttttaaaat	4740
gttctagaat	gactttacgt	tgcaggtact	ggttaattgg	ctgttgacac	cacatctatt	4800
ttgtcttatg	attctgcagt	tttgcagtac	ttttctctat	ctgattcagc	catttctgcc	4860
agagggaaaa	ggtcggcaga	aaagatgtat	tgagtgaata	gttaaggata	ggatctttgt	4920
ccaaaaattt	cagaaagatt	gagcaaactc	gacgtattca	ttgagtgagt	ttctgtgttt	4980
tcaaagggtg	aggagaaatt	tgtgctggaa	gtttttaagc	ctccgttttc	ttggaaatca	5040
gtctgtaaca	ctggcaagtc	ttaagatagt	cccatttaga	ctttgcagat	gctgaacctg	5100
gctctgtaac	gctgggaagt	cttaagatag	tcctgttttag	actttgcaaa	ccctgtacct	5160
ggcttttctc	ggagattcgg	gatgctggct	cctgcaggca	gggcgtgtgg	gagcctcgtc	5220
agaaagtttt	agaggtttcc	agcagaagca	gaatgaagat	ggctctccctg	gccttttccct	5280
taatttctcaa	ttttgattga	gggtgcacaag	ttgactttta	aagccaacgc	ttaagatact	5340
gattgacatc	ttcaagggag	aatgctccca	ggaggggctg	aagaagccat	agttggaagt	5400
ggaagggtact	cgtcagtggt	ctccacaaac	ctttttactc	tgttgtctca	gccgcactgg	5460
ggcggaggcg	gtcaaggggtg	agaagtaccg	acactcaagt	gcaaactgcc	acgtcgttgg	5520
cccatcccat	cagtgggcag	ctggctgacg	ccattcactg	gacgggtccct	gaacacctag	5580
gaatgcacac	accgtgcttc	tcagacactg	gagacgcaaa	ggcaggagga	tgacgtccgg	5640
tgagaggaca	cgatctttac	ctgcacaatc	agactgtaag	cccagcagag	aaccccaggg	5700
gcgcctgggt	acttctcgga	ggtcatctta	gttgtggtgg	ggaagacaaa	gaaataagca	5760
aacaagaaac	tagagttact	atacaagaaa	ctctcctgag	tttgtaaacc	ttaagcataa	5820
ggattcagtt	gacctttttc	ttggttcatc	aatctggaaa	gaacttacat	aaagcgccat	5880
tgacactgtc	acctgggagc	tccatggggc	gtaagtcttt	gacagccaat	ttaatttgag	5940
gtcagagggc	cttgagggtac	acagtcagca	ctgtttgaac	acttttcctg	aaagcaaaac	6000
tcacagctcc	ctgcgccctc	tgacaacact	agctatttct	gccagagtaa	gaacttctat	6060
tactatttta	ttattgttca	tatgtctttt	gatgatgggt	gtgtgacagg	gggaagcagg	6120
atctatttgg	tttcttcccc	tccccccacc	ccttcctttt	tgtctctctt	tttttttctc	6180
taagaaaatc	accagactag	tttttccatc	ttgagtaatt	tcttatgtgg	gacagttttg	6240
atcctcattt	tgaaagcatg	cgtgtgcaca	tgtgtgttgc	ctgtgggtgcc	aggtgagaca	6300
gggtggcacta	actccagctg	cttggaaggc	atcccaaggg	cgcatcttaa	agttggagca	6360
gacctccctt	ttccagcccc	tggggccatt	agaccacgtg	ctggaactag	cattgtaaaa	6420
ttcccatccc	agttccactc	ccctgaagtg	aaaccctttt	ttttttgtga	cagtaaactc	6480
taaaaatcat	tgtctcttta	tgaacatttc	ctcagtttct	tctctgctga	aaatgtaagc	6540
catgctactt	tttaatgtat	tttgaatttt	gtgctcattg	gaaattgata	tgctaatgcc	6600
tccccacccc	cccgccagac	ttttcttttt	atactttgtc	ttgttttttac	tggggtaggc	6660

tgggcatgcg tgcgtgcctt tagggcagca ttttaaacct ttgccaaaat tgcaaaggg	6720
acatgtacat tcttctgctc catcctactt aaacacctat cagctatctt tatctttaac	6780
cttttctgta tgtttgaagt gtgtgggggg tgtgtgtgtg tgtgaaagag cgagagaatg	6840
atgtcatcta aagttttttg aagaattatt tggttttcat tgcattaaaa ttctatcact	6900
cccagctttg ttttcattta aaaaaatata caaagagctt tgtaaataca acacatttta	6960
tttctcccc ttcttttaat gtacagcttt tttgccactt atatatactt aaaatattcc	7020
catgaattat gtccagttct tcttggaata aaatttggtt ttgaatgaac ctgcaaagca	7080
tcctgcagcg tgagcagctc ctccacctgg agctccgaag catcttctca ggccaaagcg	7140
gcattaccg tgaatctgtc ttctccgcca cagcatgggt tgaggcgag tctgttaata	7200
tagctgggccc atgtcagtga ctgttgtgtt tgtgggggtca ggtggggggc atggtatttg	7260
caaaaaaaaaac aaattatggc taatttatta tttgttgca gtgggggttaa ctgtaaactc	7320
atgtaagagt ctgtgatttc ctcatgggt gatctctctc tctgtaatcc tcattgcaaa	7380
ttttcaccag gacagcgttt tttgattaga ggggagctct ggcacagtat gctttaattt	7440
agcaggaact tccagatgat ttaaattctc gatgctgtga tgacacacat atgatctttc	7500
gtgtttctga gcgactctac tttcattggt tgccagcgtg gctcgttgct gttgcccaat	7560
aaagcttggtg tacgttc	7577

<210> 105

<211> 1672

<212> DNA

<213> NM\_004503.2| Homo sapiens homeo box C6 (HOXC6), transcript variant 1, mRNA

<400> 105	
ttttgtctgt cctggattgg agccgtccct ataaccatct agttccgagt acaaactgga	60
gacagaaata aatattaaag aaatcataga ccgaccaggt aaaggcaaag ggatgaattc	120
ctacttctact aacccttcct tatcctgcc cctcgccggg ggccaggagc tcctcccaa	180
cgtcgccctc aattccaccg cctatgatcc agtgaggcat ttctcgacct atggagcggc	240
cgttgcccag aaccggatct actcgactcc cttttattcg ccacaggaga atgtcgtgtt	300
cagttccagc cgggggcccgt atgactatgg atctaattcc ttttaccagg agaaagacat	360
gctctcaaac tgacagacaa acaccttagg acataacaca cagacctcaa tcgctcagga	420
ttttagttct gagcagggca ggactgcgcc ccaggaccag aaagccagta tccagattta	480
cccctggatg cagcgaatga attcgcacag tggggtcggc tacggagcgg accggaggcg	540
cggccgcccag atctactcgc ggtaccagac cctggaactg gagaaggaat ttcacttcaa	600
tcgctaccta acgcggcgcc ggcgcacga gatcgccaac gcgctttgcc tgaccgagcg	660

acagatcaaa atctgggttcc agaaccgccg gatgaagtgg aaaaaagaat ctaatctcac	720
atccactctc tcggggggcg gcggaggggc caccgccgac agcctgggcg gaaaagagga	780
aaagcgggaa gagacagaag aggagaagca gaaagagtga ccaggactgt ccctgccacc	840
cctctctccc ttctctccctc gctccccacc aactctcccc taatcacaca ctctgtatatt	900
atcactggca caattgatgt gttttgattc cctaaaacaa aattagggag tcaaactgtg	960
acctgaaagt cagctctgga cccctccct caccgcacaa ctctctttca ccacgcgcct	1020
cctcctctc gctcccttgc tagctcgttc tcggcttgtc tacaggccct tttccccgtc	1080
caggccttgg gggctcggac cctgaactca gactctacag attgccctcc aagtgaggac	1140
ttggctcccc cactccttcg acgccccac cccgcctccc cgtgcagaga gccggctcct	1200
gggcctgctg gggcctctgc tccaggccct caggggcccg cctggcagcc ggggagggcc	1260
ggaggcccaa ggagggcgcg ccttgcccc acaccaacc ccagggcctc cccgcagtcc	1320
ctgcctagcc cctctgcccc agcaaagtgc cagcccaggc aaattgtatt taaagaatcc	1380
tgggggtcat tatggcattt tacaaactgt gaccgtttct gtgtgaagat ttttagctgt	1440
atttgtggtc tctgtattta tatttatgtt tagcaccgtc agtgttccta tccaatttca	1500
aaaaaggaaa aaaaagaggg aaaattacaa aaagagagaa aaaaagtga tgacgtttgt	1560
ttagccagta ggagaaaata aataaataaa taaatccctt cgtgttacct tcctgtataa	1620
atccaacctc tgggtccgtt ctcgaatatt taataaaact gatattattt tt	1672

<210> 106

<211> 3394

<212> DNA

<213> NM\_004764.2| Homo sapiens piwi-like 1 (Drosophila) (PIWIL1), mRNA

<400> 106	
gttggcctcg ggctgaggtg caaggaccag gactagggcg agggcagcgg tccaagaaat	60
agaaaacaat gactgggaga gcccagacca gagccagagg aagggcccg cgtcaggaga	120
cagcgcagct ggtgggctcc actgccagtc agcaacctgg ttatattcag cctaggcctc	180
agccgccacc agcagagggg gaattatttg gccgtggacg gcagagagga acagcaggag	240
gaacagccaa gtcacaagga ctccagatat ctgctggatt tcaggagtta tcgttagcag	300
agagaggagg tcgtcgtaga gatattcatg atcttggtgt gaatacaagg cagaacctag	360
accatgttaa agaatacaaa acagggtctt caggcattat agtaaggtta agcactaacc	420
atttccggct gacatcccgt cccagtgagg ccttatatca gtatcacatt gactataacc	480
cactgatgga agccagaaga ctccgttcag ctcttctttt tcaacacgaa gatctaattg	540
gaaagtgtca tgcttttgat ggaacgatat tatttttacc taaaagacta cagcaaaagg	600

ttactgaagt	ttttagtaag	acccggaatg	gagaggatgt	gaggataacg	atcactttaa	660
caaatgaact	tccacctaca	tcaccaactt	gtttgcagtt	ctataatatt	attttcagga	720
ggcttttgaa	aatcatgaat	ttgcaacaaa	ttggacgaaa	ttattataac	ccaaatgacc	780
caattgatat	tccaagtcac	aggttggtga	tttggcctgg	cttcactact	tccatccttc	840
agtatgaaaa	cagcatcatg	ctctgcactg	acgttagcca	taaagtcctt	cgaagtgaga	900
ctgttttgga	tttcatgttc	aactttttatc	atcagacaga	agaacataaa	tttcaagaac	960
aagttttcaa	agaactaata	ggtttagttg	ttcttaccaa	gtataacaat	aagacataca	1020
gagtggatga	tattgactgg	gaccagaatc	ccaagagcac	ctttaagaaa	gccgacggct	1080
ctgaagtcag	cttcttagaa	tactacagga	agcaatacaa	ccaagagatc	accgacttga	1140
agcagcctgt	cttggtcagc	cagcccaaga	gaaggcgggg	ccctgggggg	acactgccag	1200
ggcctgccat	gctcattcct	gagctctgct	atcttacagg	tctaactgat	aaaatgcgta	1260
atgattttta	cgtgatgaaa	gacttagccg	ttcatacaag	actaactcca	gagcaaaggc	1320
agcgtgaagt	gggacgactc	attgattaca	ttcataaaaa	cgataatgtt	caaagggagc	1380
ttcgagactg	gggtttgagc	tttgattcca	acttactgtc	cttctcagga	agaattttgc	1440
aaacagaaaa	gattcaccaa	ggtggaaaaa	catttgatta	caatccacaa	tttgcagatt	1500
ggtccaaaaga	aacaagaggt	gcaccattaa	ttagtgttaa	tccactagat	aactggctgt	1560
tgatctatac	gcgaagaaat	tatgaagcag	ccaattcatt	gatacaaaat	ctatttaaag	1620
ttacaccagc	catgggcatg	caaatgaaaa	aagcaataat	gattgaagtg	gatgacagaa	1680
ctgaagccta	cttaagagtc	ttacagcaaa	aggtcacagc	agacaccag	atagttgtct	1740
gtctgttgtc	aagtaatcgg	aaggacaaat	acgatgctat	taaaaaatac	ccgtgtacag	1800
attgccctac	cccaagtcag	tgtgtggtgg	cccgaacctt	aggcaaacag	caaactgtca	1860
tggccattgc	tacaaagatt	gccctacaga	tgaactgcaa	gatgggagga	gagctctgga	1920
gggtggacat	ccccctgaag	ctcgtgatga	tcgttgcat	cgattgttac	catgacatga	1980
cagctgggcg	gaggtcaatc	gcaggatttg	ttgccagcat	caatgaaggg	atgaccgcgt	2040
ggttctcacg	ctgcatattt	caggatagag	gacaggagct	ggtagatggg	ctcaaagtct	2100
gcctgcaagc	ggctctgagg	gcttggaata	gctgcaatga	gtacatgcc	agccggatca	2160
tcgtgtaccg	cgatggcgta	ggagacggcc	agctgaaaac	actggtgaac	tacgaagtgc	2220
cacagttttt	ggattgtcta	aaatccattg	gtagagggtta	caaccctaga	ctaacggtaa	2280
ttgtggtgaa	gaaaagagtg	aacaccagat	tttttgctca	gtctggagga	agacttcaga	2340
atccacttcc	tggaacagtt	attgatgtag	aggttaccag	accagaatgg	tatgactttt	2400
ttatcgtgag	ccaggctgtg	agaagtggta	gtgtttctcc	cacacattac	aatgtcatct	2460
atgacaacag	cggcctgaag	ccagaccaca	tacagcgctt	gacctacaag	ctgtgccaca	2520
tctattacaa	ctggccaggt	gtcattcgtg	ttcctgctcc	ttgccagtac	gcccacaagc	2580
tggcttttct	tgttggccag	agtattcaca	gagagccaaa	tctgtcactg	tcaaaccgcc	2640

tttactacct ctaacctgca gaagacgatg cagccgcttt tctttttgaa atgacttttg	2700
gattttttta agctttttatt tacttttttt ttaactgtta tctttctgga tgaaacttgg	2760
gaaggggatt aggagatcta gcatttttatt tctagcattg ctattcaccg gcttccttat	2820
tttatacgta aaaattaaga ttttatattt tatcttcttg tttctcatag atattttgtg	2880
agcatttttt tgttttattt gaagaaatgt ggataagata cttggtagta taaaacagac	2940
tctctgagag tatttgaaat gtgtttggag atttacttaa acgtactttc aggagtgagc	3000
aagtcctact tataaaccta tattaacttt atttttgaga tacctgtttt gaattttaaag	3060
gagataagag gcgtaaagta ggatgctcac tacaaccata ggtgggggtt cagctcatat	3120
cttaagata aaaggtaacta ttatataacc tatacacaag atacaggaga aaatatgctt	3180
gatttttatt tggcaggggg gctagggttg atgggagtaa aaaaaacatt gaaaaatttt	3240
aaattgtcca aagaaacatt ttaagactct ttaacaaaaa aggccatgag taaatctcta	3300
tattaacatt actattttatt ttgttttgga actgggacat gattctattt gttataaaat	3360
aaaattgatg tgattgtcaa aaaaaaaaaa aaaa	3394

<210> 107

<211> 2524

<212> DNA

<213> NM\_000249.2| Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), mRNA

<400> 107	
attggctgaa ggcacttccg ttgagcatct agacgtttcc ttggctcttc tggcgccaaa	60
atgtcgttcg tggcaggggt tattcggcgg ctggacgaga cagtggtgaa ccgcatcgcg	120
gcgggggaag ttatccagcg gccagctaata gctatcaaag agatgattga gaactgttta	180
gatgcaaaat ccacaagtat tcaagtgatt gttaaagagg gaggcctgaa gttgattcag	240
atccaagaca atggcaccgg gatcaggaaa gaagatctgg atattgtatg tgaaagggtc	300
actactagta aactgcagtc ctttgaggat ttagccagta tttctaccta tggctttcga	360
ggtgaggctt tggccagcat aagccatgtg gctcatgtta ctattacaac gaaaacagct	420
gatggaaagt gtgcatacag agcaagttac tcagatggaa aactgaaagc ccctcctaaa	480
ccatgtgctg gcaatcaagg gaccagatc acgggtggagg acctttttta caacatagcc	540
acgaggagaa aagcttttaa aaatccaagt gaagaatatg ggaaaatttt ggaagttgtt	600
ggcaggtatt cagtacacaa tgcaggcatt agtttctcag ttaaaaaaca aggagagaca	660
gtagctgatg ttaggacact acccaatgcc tcaaccgtgg acaatatctg ctccatcttt	720
ggaaatgctg ttagtcgaga actgatagaa attggatgtg aggataaaac cctagccttc	780
aaaatgaatg gttacatatc caatgcaaac tactcagtga agaagtgcac cttcttactc	840

ttcatcaacc atcgtctggt agaatcaact tccttgagaa aagccataga aacagtgtat	900
gcagcctatt tgcccaaaaa cacacaccca ttctgtacc tcagtttaga aatcagtcctc	960
cagaatgtgg atgttaatgt gcaccccaca aagcatgaag ttcacttcct gcacgaggag	1020
agcatcctgg agcgggtgca gcagcacatc gagagcaagc tcctgggctc caattcctcc	1080
aggatgtact tcacccagac tttgctacca ggacttgctg gcccctctgg ggagatgggt	1140
aaatccacaa caagtctgac ctctgtcttct acttctggaa gtagtgataa ggtctatgcc	1200
caccagatgg ttctgtacaga ttcccgggaa cagaagcttg atgcatttct gcagcctctg	1260
agcaaacccc tgtccagtca gcccaggcc attgtcacag aggataagac agatatttct	1320
agtggcaggg ctaggcagca agatgaggag atgcttgaa tcccagcccc tgctgaagtg	1380
gctgccaaaa atcagagctt ggagggggat acaacaaagg ggacttcaga aatgtcagag	1440
aagagaggac ctacttccag caaccccaga aagagacatc gggaagattc tgatgtggaa	1500
atgggtggaag atgattcccc aaaggaaatg actgcagctt gtaccccccg gagaaggatc	1560
attaacctca ctagtgtttt gagtctccag gaagaaatta atgagcaggg acatgagggt	1620
ctccgggaga tgttgcataa ccactccttc gtgggctgtg tgaatcctca gtgggccttg	1680
gcacagcatc aaaccaagtt ataccttctc aacaccacca agcttagtga agaactgttc	1740
taccagatac tcatttatga ttttgccaat tttggtgttc tcagggtatc ggagccagca	1800
ccgctctttg accttgccat gcttgcccta gatagtcag agagtggctg gacagaggaa	1860
gatgggtccca aagaaggact tgctgaatac attgttgagt ttctgaagaa gaaggctgag	1920
atgcttgtag actatttctc tttggaaatt gatgaggaag ggaacctgat tggattacc	1980
cttctgattg acaactatgt gcccctttg gagggactgc ctatcttcat tcttcgacta	2040
gccactgagg tgaattggga cgaagaaaag gaatgttttg aaagcctcag taaagaatgc	2100
gctatgttct attccatccg gaagcagtag atatctgagg agtcgaccct ctcaggccag	2160
cagagtgaag tgcctggctc cattccaaac tcctggaagt ggactgtgga acacattgtc	2220
tataaagcct tgcgctcaca cattctgcct cctaaacatt tcacagaaga tggaaatatc	2280
ctgcagcttg ctaacctgcc tgatctatac aaagtctttg agaggtgtta aatatggtta	2340
tttatgcact gtgggatgtg ttcttcttct tctgtattcc gatacaaagt gttgtatcaa	2400
agtgtgatat acaaaagtga ccaacataag tgttggtagc acttaagact tatacttgcc	2460
ttctgatagt attcctttat acacagtgga ttgattataa ataaatagat gtgtcttaac	2520
ataa	2524

<210> 108

<211> 2928

<212> DNA

<213> NM\_001313.2| Homo sapiens collapsin response mediator protein 1 (CRMP1), mRNA

```

<400> 108
ccgatccggg cgggtgctggc agccggagcg gcggcgggcg ggccgagcag ccggggcagc 60
cgcgcgtggg catccacggg cgccgagcct ccgtccgtgt ctctatccct cccgggcctt 120
tgtcagcgcg cccgctggga gcggggccga gagcgccggt tccagtcaga cagccccgca 180
ggtcagcggc cggggccgagg gcgccagagg gggccatgtc gtaccagggc aagaagagca 240
tcccgcacat cacgagtgac cgactcctca tcaaagggtg acggatcatc aacgatgacc 300
aatcccttta tgctgacgtc tacctggagg atggacttat caaacaata ggagagaact 360
taatcgttcc tgggtggagt aagaccattg aagccaacgg gcggatgggt attcccggag 420
gtattgatgt caacacgtac ctgcagaagc cctcccaggg gatgactgcg gctgatgact 480
tcttccaagg gaccagggcg gcactgggtg gcgggaccac gatgatcatt gaccatgttg 540
ttcctgaacc tgggtccagc ctactgacct ctttcgagaa gtggcacgaa gcagctgaca 600
ccaaatcctg ctgtgattac tccctccacg tggacatcac aagctggtag gatggcgctt 660
gggaggagct ggaggtgctg gtgcaggaca aaggcgtcaa ttccttccaa gtctacatgg 720
cctataagga tgtctaccaa atgtccgaca gccagctcta tgaagccttt accttcctta 780
agggcctggg agctgtgac tgggtccatg cagaaaatgg agatttgata gctcaggaac 840
aaaagcggat cctggagatg ggcatacagg gtcccagagg ccatgccctg agcagacctg 900
aagagctgga ggccgaggcg gtgttccggg ccatcaccat tgcgggcccg atcaactgcc 960
ctgtgtacat caccaagggt atgagcaaga gtgcagccga catcatcgct ctggccagga 1020
agaaagggcc cctagttttt ggagagccca ttgccgccag cctggggacc gatggcaccc 1080
attactggag caagaactgg gccaaaggct cggcgttcgt gacttcccct cccctgagcc 1140
cggaccctac cagccccgac tacttgacct ccctactggc ctgtggggac ttgcagggtca 1200
caggcagcgg ccaactgtcc tacagcactg cccagaaggc ggtgggcaag gacaacttta 1260
ccctgatccc cgaggggtgtc aacgggatag aggagcggat gacggtcgtc tgggacaagg 1320
cgggtggctac tggcaaaatg gatgagaacc agtttgtcgc tgtcaccagc accaatgcag 1380
ccaagatctt taacctgtac ccaaggaaag ggcggattgc cgtgggctcg gatgccgacg 1440
tgggtcatctg ggacccccgac aagttgaaga ccataacagc caaaagtcac aagtcggcgg 1500
tggagtacaa catcttcgag ggtatggagt gccacggctc cccactagtg gtcacagcc 1560
agggcaagat cgtctttgaa gacggaaaca tcaacgtcaa caagggcatg ggccgcttca 1620
ttccgcggaa ggcgttcccc gagcacctgt accagcgcgt caaaatcagg aataagggtt 1680
ttggattgca aggggtttcc aggggcatgt atgacggctc tgtgtacgag gtaccagcta 1740
caccctaaata tgcaactccc gctccttcag ccaaatcttc gccttctaaa caccagcccc 1800
caccatcag aaacctccac cagtccaact tcagcttatc aggtgcccag atagatgaca 1860

```

acaatcccag	gcgcaccggc	caccgcatcg	tggcgccccc	tggtggccgc	tccaacatca	1920
ccagcctcgg	ttgaacgtgg	atgcgcggag	gagctagcct	gaaggattct	gggaatcatg	1980
tccatccctt	ttcctgtcag	tgtttttgaa	acccacagtt	ttagttgggtg	ctgatggagg	2040
gagggggaag	tcgaaggatg	ctctttccct	tttctgttta	ggaagaagtg	gtactagtgt	2100
ggtgtgtttg	cttggaatt	ccttgcccca	cagttgtgtt	catgctgaat	ccacctcgga	2160
gcatggtgtt	ttcattcccc	cttcctagt	aaccacaggt	tttagcattg	tcttgttctg	2220
tcccttccac	ttctaactcc	actggctcca	tgattctctg	agtgggtggtt	cctttgcacc	2280
ctgtagatgt	tctaggatag	ttgatgcatg	ttactaaatt	acgtatgcaa	gtctgtgagt	2340
gcgtctgagg	ggacatcgcc	aaggactgac	tgagacacga	tgccgagacc	tcaagccctg	2400
aggggcagtc	ccaaaaccct	tacagtgaag	atgtttactc	attgccccca	cctctggtcc	2460
acactagaaa	gaagctcgcc	ccacctccac	ctgtgagatc	cgatgaattct	cggaatggca	2520
ggggaagcct	tgacttaggt	tgagagaag	catcctccac	atcctgtgtc	agaaaccctg	2580
gtctccgtgg	cacttgtaac	tcaccgtgct	gtcttctggt	ctgtgtgtgt	tcttcaagcc	2640
agctctaggc	ttcaggccga	gccaggttca	cactcagaaa	gaggtctccc	catccccatt	2700
cggggctgac	gatggggggc	tgatggctgc	ccctgcgtgg	cctgagtcct	ggtccctctg	2760
aggcagttga	cggggcagtc	agatttttaa	agttttgtac	aaagttttcc	tttgtaatca	2820
ctcccatttt	tacttaacaa	ccaacttggt	gtggctctta	tttctgaatt	caaagcttgt	2880
gaaaaaataa	agaaaatgaa	ctgcccactg	aaaaaaaaaa	aaaaaaaaaa		2928

<210> 109

<211> 1609

<212> DNA

<213> NM\_002145.2| Homo sapiens homeo box B2 (HOXB2), mRNA

<400> 109						
atctccccct	ccccaaatcg	ctccattaca	taaatcgggg	ggggtgcagg	aggggggggtc	60
ccttccgata	ctccctcctg	acgccccccc	cagcagcccc	ctccccacc	attgaaagcc	120
atgaattttg	aatttgagag	ggagattggg	tttataaaca	gccagccgtc	gctcgccgag	180
tgtctgactt	ccttccccgc	tgtcttggag	acatttcaaa	cttcatcaat	caaggagtcg	240
acattaattc	ctcctcctcc	tcctttcgag	caaaccttcc	ccagcctcca	gcccggcgcc	300
tccacccttc	agagaccag	gagccaaaag	cgagccgaag	atgggcctgc	tctgccgccg	360
ccaccgcccgc	cgccactccc	cgctgcccc	ccggcccccg	agttcccttg	gatgaaagag	420
aagaaatccg	ccaagaaacc	cagccaatcc	gccacgtctc	cttctccggc	cgcctccgcc	480
gttccggcct	ccgggggtcgg	atcgctgca	gatggcctgg	gactgccgga	ggctggtggc	540



ggcggggcgc gcaggctgcg cacggcttac accaacacgc agctgctgga actggagaag	600
gaattccact ttaataagta cctgtgccgg ccacgccgcg tcgagatcgc ggccttgctg	660
gacctcaccg aaaggcaggt caaagtctgg ttccagaacc ggcgcataaa gcacaagcgg	720
cagacgcagc accgagagcc gccggatggg gagcctgcct gcccgggagc cctggaggac	780
atctgcgacc ctgccgagga acccgcggcc agcccgggcg gcccctccgc ctgcggggcg	840
gcgtgggaag cctgctgtca cccgccggag gtggtgccgg gggccttaag cgcggacccc	900
cggccttttag ccgttcgctt agagggcgca ggcgcgtcga gtcccggctg cgcgctgcgc	960
ggggccggcg ggctggagcc cgggccattg ccagaagacg tcttctcggg gcgccaggat	1020
tcacctttcc ttcccagcct caacttcttc gcggccgact cctgtctcca gctatccgga	1080
ggcctctccc ctagcctaca gggttctctc gacagcccgg tccctttttc cgaggaagag	1140
ctggattttt tcaccagtac gctctgtgcc atcgacctgc agtttcccta acctgtttcc	1200
tcctcccggt cttttcgacc cccgcgctcc ttggccgtct actggaaaaa tcgagcctct	1260
cccaccctca gtcgcataga cttatgtgtt ttgctaaaat tcaggtatta ctgaattagc	1320
gtttaatcca cttcctttct tcttcttcta aaatattggg cactcgggta tcttttaaaa	1380
ttcacacaga aaaattccgt ttggtagact ccttccaatg aaatctcagg aataattaa	1440
ctctaggggg actttcttaa aaataactag agggacctat tttcctcttt ttatgtttt	1500
agactgtaga ttatttatta aaattcttta ataataaggaa aaggggaaag tattttattgt	1560
acattatttt catagattaa ataaatgtct ttataataacc aaaaaaaaaa	1609

<210> 110

<211> 3262

<212> DNA

<213> NM\_002860.2| Homo sapiens aldehyde dehydrogenase 18 family, member A1 (PYCS/ALDH18A1), mRNA

<400> 110	
agggggcgaa ggcggcggtg gtgaggaaga tactttggtt agtgaccaca tcgcagcatg	60
ttgagtcaag ttaccgctg tgggttccag cccttcaacc aacatcttct gccctgggtc	120
aagtgtacaa ccgtcttcag atctcattgt atccagcctt cagtcacatg acatgttcgt	180
tcttgagaca acatcccgtt tatcactgta cccctcagtc gtacacatgg aaagtccttc	240
gcccaccgca gtgagctgaa gcatgccaaag agaatcgtgg tgaagctcgg cagtgccgtg	300
gtgacccgag gggatgaatg tggcctggcc ctggggcgct tggcatctat tggtgagcag	360
gtatcagtgc tgcagaatca gggcagagag atgatgctgg tgaccagtgg agccgtagcc	420
tttgcaaac aacgcttgcg ccatgagatc cttctgtctc agagcgtgcg gcaggccctc	480
cactcggggc agaaccagct gaaagaaatg gcaattccag tcttagaggc acgagcctgt	540

gcagctgccg	gacagagtgg	gctgatggcc	ttgtatgagg	ctatgtttac	ccagtacagc	600
atctgtgctg	cccagatttt	ggtgaccaat	ttggattttcc	atgatgagca	gaagcgccgg	660
aacctcaatg	gaacacttca	tgaactcctt	agaatgaaca	ttgtcccat	tgtcaacaca	720
aatgatgctg	ttgtccccc	agctgagccc	aacagtgacc	tgcagggggg	aaatgttatt	780
agtgttaaag	ataatgatag	cctggctgcc	cgactggctg	tggaaatgaa	aactgatctc	840
ttgattgttc	tttcagatgt	agaaggcctt	tttgacagcc	ccccaggttc	agatgatgca	900
aagcttattg	atatatttta	tcccggagat	cagcagtctg	tgacatttgg	aaccaagtct	960
agagtgggaa	tgggtggcat	ggaagccaag	gtgaaagcag	ccctctgggc	tttgcaaggt	1020
ggcacttctg	ttgttattgc	caatggaacc	cacccaaagg	tgtctgggca	cgtcatcaca	1080
gacattgttg	aggggaagaa	agttggtacc	ttcttttcag	aagtaaagcc	tgcaggccct	1140
actgttgagc	agcagggaga	aatggcgcg	tctggaggaa	ggatgttggc	caccttggaa	1200
cctgagcaga	gagcagaaat	tatccatcat	ctggctgatc	tgttgacgga	ccagcgtgat	1260
gagatcctgt	tagccaacaa	aaaagacttg	gaggaggcag	aggggagact	tgcagctcct	1320
ctgctgaaac	gtttaagcct	ctccacatcc	aaattgaaca	gcctggccat	cggtctgcga	1380
cagatcgag	cctcctccca	ggacagcgtg	ggacgtgttt	tgcgccgcac	ccgaatcgcc	1440
aaaaacttgg	aactggaaca	agtgactgtc	ccaattggag	ttctgctggg	gatctttgaa	1500
tctcgctctg	actgtctacc	ccaggtggca	gctttggcta	tcgcaagtgg	caatggcttg	1560
ttactcaaag	gaggggaagga	ggctgcacac	agcaaccgga	ttctccacct	cctgaccag	1620
gaggctctct	caatccatgg	agtcaaggag	gccgtgcaac	tggatgaatac	cagagaagaa	1680
gttgaagatc	tttgccgcct	agacaaaatg	atagatctga	tcattccacg	tggctcttcc	1740
cagctgggtca	gagacatcca	gaaagctgct	aaggggattc	cagtgatggg	gcacagcgaa	1800
gggatctgtc	acatgtatgt	ggattccgag	gccagtgttg	ataaggtcac	caggctagtc	1860
agagactcta	aatgtgaata	tccagctgcc	tgtaatgctt	tggagacttt	gttaatccac	1920
cgggatctgc	tcaggacacc	attatttgac	cagatcattg	atatgctgag	agtggaacag	1980
gtaaaaattc	atgcaggccc	caaatttgcc	tcctatctga	ccttcagccc	ctccgaagtg	2040
aagtcactcc	gaactgagta	tggggacctg	gaattatgca	ttgaagtagt	ggacaacggt	2100
caggatgcc	ttgaccacat	ccacaagtat	ggcagctccc	acacggatgt	catcgtcaca	2160
gaggacgaaa	acacagcgga	gttcttcctg	cagcacgtag	acagtgcctg	tgtgttcttg	2220
aatgccagca	ctcgcttttc	tgatggttac	cgctttggac	tgggagctga	agtgggaatc	2280
agtacatcga	gaatccacgc	ccggggacca	gtaggacttg	agggactgct	tactactaag	2340
tggctgctgc	gaggggaagga	ccacgtggtc	tcagatttct	cagagcatgg	aagtttaaaa	2400
tatcttcatg	agaacctccc	tattcctcag	agaaacacca	actgaaaaga	gccaggaaaa	2460
cccgggaatt	ttccaaaagg	tcttcacgtt	aaacttgtct	tatctcagga	gagagcccgc	2520
tcttgtctcc	cagttcctgg	tagggctctgc	ctgttgga	gtgtacctgg	atgcttctgg	2580

gctccgtttg	gcaatagcaa	tcttggctga	tgtgcacagt	ctggctccca	gctcaccctt	2640
tttttttaaa	gtaagaaaat	agttgctacc	gatagggact	ttgccaagtc	caattatctt	2700
ctaggattga	aagggtgcatt	ttccccataa	aaaaggcgag	gaaaacccat	ggctgctttg	2760
tgtcacctca	gtgacttaca	gtcccccttc	gcatttagtt	ggtactagag	ccagtcatcc	2820
ttaacaaatc	ttttcgcgtt	ttatttcttt	cacatgtagt	catcttcaaa	aaggaaagat	2880
ttggaatttt	agaaaagggg	caactcttct	ttttagcatt	ctcatcagaa	agtcacaaaa	2940
atcgatggaa	tcatttccac	tgggaagatt	gaccttttgt	atttatttgt	ggggtaaatt	3000
aataagcatt	ccagatgctt	gcagcttcct	gcatccagga	gatgctgtgt	ccccgtgat	3060
gcagctggaa	cccaagctgc	agcaggagat	gcaagtttca	ggatgttccc	cactgagctg	3120
gaggaatatc	tacagcagtg	atgcttgaaa	tttttgtatg	aattattttg	tcgtcctacc	3180
cttttcctcc	aaaacaaaaa	ttagaggatt	attttaatac	tttggtattct	tccccctttt	3240
ttgagaaata	aagtttttta	tg				3262

<210> 111

<211> 2899

<212> DNA

<213> NM\_005655.1| Homo sapiens TGFB inducible early growth response (TIEG), mRNA

<400> 111	
cagacggcgc	tgagcgcggc ggcggcggga gcggcgtcga gtgtctccgt gcgcccgtct 60
gtggccaagc	agccagcagc cttagcagcca gtcagcttgc cgccggcggc caagcagcca 120
accatgctca	acttcgggtgc ctctctccag cagactgcgg aggaaagaat ggaaatgatt 180
tctgaaaggc	caaaagagag tatgtattcc tggaacaaaa ctgcagagaa aagtgatttt 240
gaagctgtag	aagcacttat gtcaatgagc tgcagttgga agtctgattt taagaaatac 300
gttgaaaaca	gacctgttac accagtatct gatttgtcag aggaagagaa tctgcttccg 360
ggaacacctg	attttcatac aatcccagca ttttgtttga ctccacctta cagtccttct 420
gactttgaac	ctctctcaagt gtcaaactcg atggcaccag cgccatctac tgtacacttc 480
aagtcactct	cagatactgc caaacctcac attgccgcac ctttcaaaga ggaagaaaag 540
agcccagtat	ctgcccccaa actccccaaa gctcaggcaa caagtgtgat tcgtcataca 600
gctgatgccc	agctatgtaa ccaccagacc tgcccaatga aagcagccag catcctcaac 660
tatcagaaca	attcttttag aagaagaacc cacctaaatg ttgaggctgc aagaaagaac 720
ataccatgtg	ccgctgtgtc accaaacaga tccaaatgtg agagaaacac agtggcagat 780
gttgatgaga	aagcaagtgc tgcactttat gacttttctg tgccttcctc agagacggtc 840
atctgcaggt	ctcagccagc ccctgtgtcc ccacaacaga agtcagtgtt ggtctctcca 900

cctgcagtat ctgcaggggg agtgccacct atgccggtca tctgccagat ggttccccctt	960
cctgccaaaca accctgttgt gacaacagtc gttcccagca ctccctcccag ccagccacca	1020
gccgtttgcc cccctgttgt gttcatgggc acacaagtcc ccaaaggcgc tgtcatgttt	1080
gtggtacccc agcccgttgt gcagagttca aagcctccgg tggtagagccc gaatggcacc	1140
agactctctc ccattgcccc tgctcctggg ttttccccctt cagcagcaaa agtcactcct	1200
cagattgatt catcaaggat aaggagtcac atctgtagcc acccaggatg tggcaagaca	1260
tacttttaaaa gttcccatct gaaggcccac acgaggacgc acacaggaga aaagcctttc	1320
agctgtagct ggaaagggtg tgaaaggagg tttgcccggt ctgatgaact gtccagacac	1380
aggcgaaccc acacgggtga gaagaaatgt gcgtgcccc a tgtgtgaccg gcggttcattg	1440
aggagtgacc atttgaccaa gcatgccccg cgccatctat cagccaagaa gctaccaaac	1500
tggcagatgg aagttagcaa gctaaatgac attgctctac ctccaacccc tgctcccaca	1560
cagtgcacaga ccggaaagtg aagagtcaga actaactttg gtctcagcgg gagccagtgg	1620
tgatgtaaaa atgcttccac tgcaagtctg tggccccaca acgtgggctt aaagcagaag	1680
ccccacagcc tggcacgaag gccccgcctg ggtaggtga ctaaaagggc ttcggccaca	1740
ggcaggtcac agaaaggcag gtttcatttc ttatcacata agagagatga gaaagctttt	1800
attcctttga atattttttg aaggtttcag atgaggtcaa cacaggtagc acagattttg	1860
aatctgtgtg catattttgtt actttacttt tgctgtttat acttgagacc aacttttcaa	1920
tgtgattctt ctaaagcact gggttcaaga atatggaggc tggaaggaaa taaacattac	1980
ggtacagaca tggagatgta aaatgagttt gtattattac aaatattgtc atctttttct	2040
agagttatct tctttattat tcctagtctt tccagtcaac atcgtggatg tagtgattaa	2100
atatacttag aactatcatt ttacactat tgtgaatatt tggaattgaa cgactgtata	2160
ttgctaagag ggcccaaaga attggaatcc tccttaattt aattgctttg aagcatagct	2220
acaatttggt tttgcatttt tgttttgaaa gtttaacaaa tgactgtatc taggcatttc	2280
attatgcttt gaactttagt ttgcctgcag tttcttgtgt agatttgaaa attgtatacc	2340
aatgtgtttt ctgtagactc taagatacac tgcactttgt ttagaaaaaa aactgaagat	2400
gaaatatata ttgtaaagaa gggatattaa gaatcttaga taacttcttg aaaaagatgg	2460
cttatgtcat cagtaaagta cctttatgtt atgaggatat aatgtgtgct ttattgaatt	2520
agaaaattag tgaccattat tcacagggtg acaaatgttg tcctgttaat ttataggagt	2580
tttttgggga tgtggaggta gttgggtaga aaaattatta gaacattcac ttttgtaa	2640
agtttttctc ttttattctg ttatatagtg gatgatatac acagtggcaa aacaaaagta	2700
cattgcttaa aatatatagt gaaaaatgtc actatatctt cccatttaac attgtttttg	2760
tatatgggt gtagatttct gacatcaaaa cttggaccct tggaaaacaa aagttttaat	2820
taaaaaaaaat ccttgtgact tacaatttgc acaatatttc ttttgttgta ctttatatct	2880

tgtttacaat aaagaattc

2899

<210> 112

<211> 3138

<212> DNA

<213> NM\_018223.1| Homo sapiens checkpoint with forkhead and ring finger domains (CHFR), mRNA

<400> 112

ctcttgacag	cggcggcggc	gcagccggtt	ccgggttcgg	cgcggggcgg	ggatgtgaat	60
cccgatggag	cggcccagag	aaggcaagca	gtcgccgccg	ccgcagccct	ggggacggct	120
cctgcgtctg	ggcgcggagg	agggcgagcc	gcacgtcctc	ctgaggaagc	gggagtggac	180
catcgggcgg	agacgagggt	gcgacctttc	cttccccagc	aataaactgg	tctctggaga	240
tcactgtaga	attgtagtgg	atgaaaaatc	aggtcagggt	acactggaag	ataccagcac	300
cagtggaaca	gtgattaaca	agctgaaggt	tgtaagaag	cagacatgcc	ctttacagac	360
tggggatgtc	atctacttgg	tgtacaggaa	gaatgaaccg	gaacacaacg	tggcatacct	420
ctatgaatct	ttaagtgaag	agcaaggcat	gacacaagaa	tcctttgaga	tgggtgccttg	480
ctgtgttgcc	caggctgggt	taaaactcct	gggatcaagt	gatcctccca	ccttggcctc	540
ccaaagtatt	gtgattacag	ggctctgggg	tgggtggcatc	tcccctaaag	gaagtgggtcc	600
ctctgtggca	agtgatgaag	tctccagctt	tgcctcagct	ctcccagaca	gaaagactgc	660
gtcctttttc	tcgttggaac	cccaggatca	ggaggatttg	gagcccgaga	agaagaaaat	720
gagaggagat	ggggaccttg	acctgaacgg	gcagttgttg	gtcgcacaa	cgcgtagaaa	780
tgcccacacc	gtccacgagg	acgtcagagc	agcggctggg	aagccagaca	agatggagga	840
gacgctgaca	tgcacatctt	gccaggacct	gctgcacgac	tgcgtgagtt	tgcagccctg	900
catgcacacg	ttctgcgcgg	cttgctactc	gggctggatg	gagcgtctgt	ccctgtgtcc	960
tacctgccgc	tgtcccgtgg	agcggatctg	taaaaaccac	atcctcaaca	acctcgtgga	1020
agcatacctc	atccagcatc	cagacaagag	tcgcagtga	gaagatgtgc	aaagtatgga	1080
tgccaggaat	aaaatcactc	aagacatgct	gcagcccaaa	gtcaggcggg	ctttttctga	1140
tgaagaaggg	agttcagagg	acctgctgga	gctgtcagac	gttgacagtg	agtcctcaga	1200
cattagccag	ccatacgtcg	tgtgccggca	gtgtcctgag	tacagaaggc	aggcggcgca	1260
gcctccccac	tgccagcac	ccgagggcga	gccaggagcc	ccacaggccc	tgggggatgc	1320
acccccccac	tccgtcagcc	tgacgacagc	agtccaggat	tacgtgtgcc	ctctgcaagg	1380
aagccacgcc	ctgtgcacct	gctgcttcca	gcccattgcc	gaccggagag	cggagcgcca	1440
gcaggaccgg	cgtgtcgccc	ctcagcagtg	tgcggctctg	ctgcagcctt	tctgccacct	1500
gtactggggc	tgcacccgga	ccggctgcta	cggctgcctg	gccccgtttt	gtgagctcaa	1560

```

cctgggtgac aagtgtctgg acggcgtgct gaacaacaac agctacgagt cagacatcct 1620
gaagaattac ctggcaacca gaggtttgac atggaaaaaac atgttgaccg agagccctcg 1680
ggctctccag cggggagtggt ttctgctgtc tgattacaga gtcacgggag acaccgttct 1740
gtgttactgc tgtggcctgc gcagcttccg tgagctgacc tatcagtatc agcagaacat 1800
tcctgcttcc gagttgccag tggccgtaac atcccgtcct gactgctact ggggccgtaa 1860
ctgccgcact caggtgaaaag ctccaccacgc catgaaattc aatcatatct gtgaacagac 1920
aaggttcaaa aactaagcat ccagaggccc tgagcagctt tcagcactgg aggtgaagag 1980
agcgtgtttt taaaatacag aggcaagcac gtcaagggtg ttacacagcc ccctgagggg 2040
agggacgcag ggtctccgac aggtgctctg gggtgactct tctgtggagc ttaccctct 2100
gagtgaagacc ctcccagag ccccgggggc cgcagccgc cctcctggtg agcgtgggc 2160
agggctcgtg gtggcatcag cagcagagac gaagcctttc tgtaacatgc ggccgtcctg 2220
ccgagagggg cagttttgct cttttgtaca ttttccgaaa ctacagttaa agcgggaagtc 2280
tgttttcagg aaaagtttca agggagaagg gcaagtttat caaaaacatt gtttcaggag 2340
aagggaagcat aagtttacag cctacaggac gtacacaata tcctgctgct gggaaaacca 2400
cagcatttta tctatttttt attttaatag gtttggtgct tatcttctaa taagatttaa 2460
atgtcacaaa ctgtagcaca aataatataa tttataattt acaaattgac taaaattggg 2520
tatagtatgg tatttgaaag aataagcata tgcttctgtt tattaaaaaa agaaaccttc 2580
caatgtccaa aactgctaac cctcgacgtg gccgccaagt tagtcgctcc ttgctaaccg 2640
gtgagtgacc gcggccccga gcctggggct ggacgcaggt cccaggacat gctgctccct 2700
tgtgtgagtg accgcggccc cgagcctggg gctggacgca ggtcccagga cgtgctgctc 2760
ccttgtgtga gtgaccacgg cccaagccc agggctggag gcaggtcca ggacgcgccg 2820
ctccctcatg ctgcccgggc ctttctcca agaccctaca gagcctgagg ggcaccttg 2880
cttccgcctg tgctagcttt gccatgtcat ctggaataat acttgaaatt ttgattcttg 2940
gaaaaaaaaa tttcttatct tttgttgaaa tcacctgtta tccttgtttg taaactgata 3000
acttttttgc ttcttctcag gaatacagtt ttcaactgtt gtcttgctct tgatagaaac 3060
tgagaagcag caatctgtat ttgtggagga aagtcctctc ttttgcata tctaataaat 3120
gagccgcggt tgctcctc 3138

```

<210> 113

<211> 2466

<212> DNA

<213> NM\_024645.1| Homo sapiens hypothetical protein FLJ13842 (FLJ13842), mRNA

<400> 113

```

agttgggtccg agctgccgaa aggtctgggt gcagagacag gaacgtgtaa tcctcagcgt 60

```

gctccagccc	acagcttcgc	tctactgctc	ggcagggcag	ctggcctctg	ggcaccggcg	120
gccccctctgc	ctcgcgga	agcctgatga	agtcctccga	tattgatcag	gattttattca	180
cagacagtta	ctgcaaggtg	tgcagtgcac	agctgatctc	cgaatcgag	cgtgtggccc	240
actacgagag	tcgaaaacat	gcaagcaaag	tccgactgta	ttacatgctt	caccccaggg	300
atggaggggtg	tcctgccaag	aggctccggg	cagaaaatgg	aagtgatgcc	gacatggtgg	360
ataagaacaa	gtgctgcaca	ctctgcaaca	tgtcattcac	ttcagcgggtg	gtggccgatt	420
cccattatca	aggcaaaatc	cacgccaaga	ggttaaaact	cttgctagga	gagaagaccc	480
cattaaagac	cacagcaaca	cccctgagcc	cacttaagcc	cccacggatg	gacactgctc	540
cgggtggtcgc	atctccctat	caaagaagag	attcagacag	atactgtggg	ctctgtgcag	600
cctggtttaa	taaccctctg	atggcccagc	aacattatga	tggcaagaaa	cacaaaaaga	660
atgcggaag	agttgctttg	ttagaacaac	tggggacaac	cctggatatg	ggggaactga	720
gaggtctgag	gcgcaattac	agatgtacca	tctgcagtgt	ctccctaaac	tcaatagaac	780
agtatcatgc	ccatctgaaa	ggatctaaac	accagaccaa	cctgaagaat	aagtagtgaa	840
agcatcaatc	aagacataag	aacaaaacat	tagcatttct	ctgccgtgga	gaattgctta	900
tcaaccacca	gaggaggctt	ctttcttgaa	caataaacat	ttcttataag	gattcacaga	960
ttcacatacg	actgatcttg	atTTTTggaa	atgaatgagg	tttctTTTTt	cttttccctt	1020
TTTTtaattt	tggggtaagt	tatgatattt	ggatggattt	ttaaattctt	tcctgataac	1080
atatttagca	catgttctaa	attataatcc	tatagcaaac	agttggagca	ttattcaaac	1140
tgaagtgga	aaaattttaa	tttccaattt	attctagatt	tcctcagagc	ataattattc	1200
tgttaaattc	tcaatgagtg	tgatgtaaac	cacctctatc	cagaaatata	cattcttttc	1260
tcatcatgtt	ggacacagtt	gagggtgaca	tgcacagaac	tggaacagat	cactattagt	1320
ggaaaatacc	aaatggacaa	ataaatacca	gtcgttttct	ccgttctcca	agcacaggag	1380
ccaggttttac	catctgaaca	atgaagacga	agggagtaaa	taaaggaaga	attctcatct	1440
TTTTtcctga	tcattcaaag	aacagtttct	caaggttaag	ccaagtcctc	cttgcaagtt	1500
gccaaataat	agcttaggaa	aagaattagt	ctgcctgcat	gatgatcttc	ttaggcaaaa	1560
acgtcttcac	agcccttgac	cttggtgaat	TTTTttcccc	aaaagcatcc	aaaagaagaa	1620
ttataaacc	cagaacgaga	tggaaataaa	caagtatttt	TTTTttatga	tgtttggcct	1680
gaactgtggg	ctttaattgg	gggatactga	tcgtttggaa	agaagtgaga	aaattctgaa	1740
gaaatggcgg	ccttgggcta	ggcgggggtcc	cctatttctt	ctgtttctca	ctgaagtcct	1800
actgctgagc	caagactcag	tcactctgga	aagagcatga	ccgataaaga	aaacagttcc	1860
tttctgatgg	ggagcgtctg	agtgcagatc	atgaggctct	ttctctaggt	ttaattcttt	1920
tccatggtga	ccggacttgg	tgtctttag	cctgggttac	aagtgggacg	ttgagcttct	1980
actgacgatg	ccctgcatgg	accagctggg	atctggctgg	ggctgccctg	tgtccctaac	2040

gaccataggc aatccatctt cttgtgtcag caatttctgg acaccactg tttccacca	2100
agagctgagg tggcaacaac tcagtgagca ataaacaaaa tgacacagaa atgcacagt	2160
ttgttatgaa ggagcctgtt tacctgtgtt caaaatctgg caccattccc ttgagcaggg	2220
cccgtctagg agggaccagg tctgccagtt tctgtgcctg cagagagacg aagccccacg	2280
agccacaccc tactctacaa gaggaaaggg ggttggatgg gaagaatcta ttttgctgtt	2340
ttggaagca cacagccgac ctacaaacct cctgtgatgg tgtttcttcg gatgtgtaaa	2400
ataaggcttt atttgtcaat tccgctgtaa aataagcatt gtccgagtaa aaacagcagc	2460
aacaac	2466

<210> 114

<211> 3658

<212> DNA

<213> NM\_025195.2| Homo sapiens tribbles homolog 1 (Drosophila) (TRIB1), mRNA

<400> 114	
aaggaggagt gcactggccg ggatcgggtgc agcgctcaca ctcactcaca gtcactctct	60
ctgagcgcggt ctcgctcgct ctcatcacag cccggagccc aggagcgctc aggatcccga	120
gcgccgcgaa aaagttcccc cggtttttgc tggagactca tcgttttggg aagtgcattt	180
gcttcgtggc tccgccgagc ctgctgaatc ctgtcctcgc ggcacgggac cccgggatcg	240
ctgaccgctg ccgcccgcgc ctctgcctcc cggactatcg gcagcctcgg caacaatagt	300
ggcgcccgcc cccagcgagg ctccgggagc ccttgctgc ggggggtccgg ggactcgagc	360
cggcctccgc ctcccggacg cacagccagc gtggtccccg cgtgcaacgc gagcgccggg	420
gagtggctcc tgctttgccc ctctgtgggg ccgagccaag accagtctgc aaactccatc	480
ccgcccggctg gaagaagtcg cggagccggc accaaacccg cagcgtcttc ccgcgcggat	540
cccgggactt aaaaagccgg ggccaccccc gccaggagc ggatgcgggt cgggtccggtg	600
cgtcttgcca tgagcggcgc ctgcgagccc cgcggccccg ccctgctctt cccagccacc	660
cgaggcgtcc cggccaaacg cctgctggac gccgacgacg cggcggctgt ggcggccaag	720
tgcccgcgcc tctccgagtg ctccagcccc ccggactacc tcagcccccc cggctcgccc	780
tgcagcccgc agcccccgcc tgccgctccg ggggcccggc gaggtccgg gagcgcgccg	840
gggcccagcc gcatcgccga ctacctgctg ctgcccctag ccgagcgca gcatgtgtcc	900
cgggcgctgt gcatccacac tggacgcgag ctgcgctgca aggtgtttcc cattaaacac	960
taccaggaca aaatcaggcc ttacatccag ctgccatcgc acagcaacat tactggcatt	1020
gtggaagtga tccttgggga aaccaaggcc tatgtcttct ttgagaagga ctttggggac	1080
atgcactcct atgtgcgaag ccggaagagg ctgcgggaag aggaagccgc cgggctcttc	1140
aagcagattg tctccgccgt cgcccactgc caccagtcag ccatcgtgct gggggacctg	1200



aagcttagga	agttcgtctt	ctccacggag	gagagaaccc	agcttagact	agaaagtcta	1260
gaagacacac	acataatgaa	gggggaagat	gatgctttgt	cagacaaaca	tggctgcccc	1320
gcctacgtga	gccctgagat	cctcaacacc	actgggacct	actccggaaa	ggctgcggac	1380
gtttggagcc	tgggggtgat	gctctacacc	cttctggttg	gacgataccc	cttccatgac	1440
tcagacccca	gtgccctttt	ctccaaaatt	cggcgtggac	agttctgcat	tcctgagcac	1500
atttccccc	aagccaggtg	cctcattcgc	agcctcttga	gacgggagcc	ctccgagaga	1560
ctcactgccc	ccgagatcct	actgcacccc	tggtttgagt	ccgtcttgga	acccgggtac	1620
atcgactcag	aaataggaac	ttcagaccag	attgtttccag	agtaccagga	ggacagtgc	1680
attagttcct	tcttctgcta	atccccaaaa	cctcagaaac	ctcataattc	ttaacacctg	1740
gcatttccat	ttctaaagat	ggacaggccc	tttggcgtgg	taccaaccag	ataatgactg	1800
catcaggatg	aaagctgctg	aactcggcat	ggcgcctcct	cttctctggt	gggatgagtg	1860
actttattga	tttgagcagc	atatgctgtg	attggctgcc	ctgcaaattt	gtttccctta	1920
aggaaccctc	accaactatc	tctgctggat	ttgggagttc	cgcattcttt	gtggagggca	1980
gagtatggac	atcttacacc	cgggtggtcaa	gtgtgtaata	aacttgagca	ttcgaatggg	2040
agaaaaagca	aatcgacaaa	tgacatattt	tgagtaataa	ccgtattttt	cacagggtga	2100
caaattgggc	caataaatct	gccatctttg	aactcatctt	tgggtggctag	actgctacgg	2160
cagcttctct	gatgggaaa	gttccttttt	ggcttaacac	tcaccctttc	ttcacactca	2220
catttaccaa	tgactctgct	ccgttttttg	agcagactgt	tttaagttgc	tcaggagcct	2280
gatggaacca	tgaaccgaga	ctcttctctg	tttcttgcca	agacctcatc	tgactaatg	2340
ccttctccct	gaccttgaca	cttccccctt	tagctataaa	agcacttacc	agccgaacgt	2400
ggaacagtat	cacaaaagat	tccatctccc	aacgatttca	gaactctgag	ctcagagaga	2460
ctccagattt	taaaaaataa	tttgagtgtc	tggaaactat	tagcttttta	agttccttcc	2520
aaatatgtta	gtacctaccc	tttacttttt	ccccaaagacc	atctcagggt	ggagcattct	2580
gtctaagaga	agaaagataa	ggaggctccc	accacctct	cccaagagca	gacattaaac	2640
atctttgtgc	tttgaagaga	gtgaattttg	gatagtcttg	tgattctcag	actaacttcc	2700
agaattatac	tttaaccctt	cccagatatg	gtccgccttt	ggcattgtgt	gtacatctgc	2760
agttttgcat	gggtgggttg	taatatttca	aatgtgtggt	ttatgaatac	gtctgtataa	2820
tcggcttctg	gagtgaaaaca	gcaaacccca	aatcttcaaa	gttggaaagga	actttaaaaa	2880
tcacccggtc	caatctcttt	cctctttctg	ccacctcca	aggcagaaat	cccctcttca	2940
gcttcttttg	taggtgggaa	tccagcctct	gttagatatg	tccagagatg	gaaactcact	3000
cccctacaaa	agatggagct	taatggagaa	attgcaactt	tcattaaaaa	acaaattcag	3060
atgaaatata	agtaactgtc	ttggacagtg	ctgaaatcag	gtggttaaac	gggtaaacaa	3120
aatatactgt	attttgagaa	atggcacaaa	aacaggcagt	catctttaag	ggctatgcct	3180

```

aggcaaacta ctaacatgca ttgtgagaat gccgtgtata cctcacgtac tgtgtacttt 3240
gtacatatat tttacctttt atacctatgt tcgattttgt tttgttttgt tttgttctgg 3300
ctttgaggct tgttttgttg tctgtgtctg tctgaataac ctgctgtctt aaaaccacgt 3360
gaaatgtgaa tgattattgg caatattacc ttgacagaat catgggactt tgagaagagg 3420
gaggacagag gcctctgtcg cactaacgct ctctgtggtt ctctgactgtt gtatctgtga 3480
tacattatcc gactaaggac tctgggctgg cagggccttc tgccgggaaa gctagaaaca 3540
ctaggttctt cctgtacata cgtgtatata tgtgaacagt gagatggccg tttctgactt 3600
gtagagaaat ttttaataaac ctggtttcgt aaaaaaaaaa aaaaaaaaaa aaaaaaaa 3658

```

<210> 115

<211> 4624

<212> DNA

<213> NM\_033331.1| Homo sapiens CDC14 cell division cycle 14 homolog B (S. cerevisiae) (CDC14B), transcript variant 2, mRNA

```

<400> 115
cacggaacag ccctcctggg gtccccacga gccgcgtcct gctgtgcccc ggcgcctacg 60
cagcagcggc cgcgcccgcg gtgggcacgc acggttacct cgggcagctc cggccgccag 120
ctgcagcccc gtcgcctcgg ccgcgccagc cggctgcggg cacctggggg cgggctgggg 180
gcgcggggcg cggcaggagg cgctgtagcg agggctgcgg cgccggtcct gcggcggccg 240
cgggaggcag cggggcaggc gctgtgggcc gggctcctcc tccggctcct gcgcgaccgc 300
ctcccgccgg gctctgccgg cgcccgccgt cccgcgacgc ccgctctgcg cccgccgccc 360
cgagcgcccc cgcggggctg gcgggagcct cggcgggcgc gcgggcgcgc ggggccatgg 420
tcgtggcccc ctgacgggcc gcggccgcct ccatgaagcg gaaaagcgag cggcggctga 480
gctgggcccgc cgcgcccccc tgctcgcggc gctgctcgtc gacctcgccg ggtgtgaaga 540
agatccgcag ctccacgcag caagaccgcg gccgcgggga ccccgaggac gacgtgtacc 600
tggacatcac cgatcgccct tgttttgcca ttctctacag cagaccaaag agtgcataaa 660
atgtacatta ttccagcata gataatgaac ttgaatatga gaacttctac gcagattttg 720
gaccactcaa tctggcaatg gtttacagat attgttgcaa gatcaataag aaattaaagt 780
ccattacaat gttaaggaag aaaattgttc attttactgg ctctgatcag agaaaacaag 840
caaatgctgc cttccttggt ggatgctaca tggttatata tttggggaga accccagaag 900
aagcatatag aatattaatc tttggagaga catcctatat tcctttcaga gatgctgcct 960
atggaagttg caatttctac attacacttc ttgactgttt tcatgcagta aagaaggcaa 1020
tgcagtatgg cttccttaat ttcaactcat ttaaccttga tgaatatgaa cactatgaaa 1080
aagcagaaaa tggagattta aattggataa taccagaccg atttattgcc ttctgtggac 1140

```

ctcattcaag agccagactt gaaagtgggt accaccaaca ttctcctgag acttatattc	1200
aatatttttaa gaatcacaaat gttactacca ttattcgtct gaataaaaagg atgtatgatg	1260
ccaaacgctt tacggatgct ggcttcgatc accatgatct tttctttgcg gatggcagca	1320
cccctactga tgccattgtc aaagaattcc tagatatctg tgaaaatgct gaggggtgcca	1380
ttgcagtaca ttgcaaagct ggccttgggt gcacgggcac tctgatagcc tgctacatca	1440
tgaagcatta caggatgaca gcagccgaga ccattgctgt ggtcaggatc tgcagacctg	1500
gctcgggtgat tgggcctcag cagcagtttt tgggtgatgaa gcaaaccaac ctctggctgg	1560
aagggggacta ttttcgtcag aagttaaagg ggcaggagaa tggacaacac agagcagcct	1620
tctccaaact tctctctggc gttgatgaca tttccataaa tgggggtcag aatcaagatc	1680
agcaagaacc cgaaccgtac agtgatgatg acgaaatcaa tggagtgaca caaggtgata	1740
gacttcgggc cttgaaaagc agaagacaat ccaaaacaaa cgctattcct ctcacagtaa	1800
ttcttcaatc cagtgttcag agctgtaaaa catctgaacc taacatttct ggcagtgcag	1860
gcattactaa aagaaccacc agatctgctt caaggaaaag cagtgttaaa agtctctcca	1920
tttcaaggac taaaacagtc ttgcgttaag taaaaacctg tgaccagagc tgaaggaaga	1980
ctctaggact gaaaactgca acagaaatta gcacaatttg aaaacaaaac aaaattgcaa	2040
aagccttagt tgctttttcc acctaagaag ttgatcaatg gagaaaatgt cactggagt	2100
ttgaataatg aactttgagt ttgggtgcaa gcaaatgact cagagaaggg tccagctctc	2160
aagctgaatg acaaacatgc tgttgtaaat ttagtctcag gtgtaaatac ccaagccctc	2220
tggtacctag ggagctggct ggtctgtggt gcatgtgtgt ccctgtgatg gcaatcattg	2280
tagttgctgg ccttcagaag aattgaggat ctgatggagg ttttttatgt atttattttc	2340
tgttcacctt gtgacctgt gtcaaaattt ataaagatac aaaaggcatt actgaaatgg	2400
tactttctgt aatttgatac tatttggtt aatcatcttc acttgactat ttgtaatact	2460
gttgtaatgt taactctgtt aagtacccaa gctgcttgtc ttccaccaa gagtgcttta	2520
ttaacaagaa tctgtgaaaa tcacatttaa acactgttgc atgttgtaag accagggtgg	2580
accttagtaa cctaaaactt gcaagagaat attaatggta gctttagaag actcaggagg	2640
agaaactgac ttcagagttg gaagatgttg caagtcgttc ctttttctgt ccttcagggg	2700
ctgaagaact gggaggctgc ccattgtttg gttgccagtc atacaaatta aaatcatatt	2760
tccttccatg aatggaagaa acacactatt ggtttttccc cttggaaaca gcaatcccaa	2820
ataatgtcgg cttacaaaaa aaaaaagtta ccactttttt agagtccttc cctgtaacat	2880
tggatttttt ttttccctta tgagatccac ctaaggccat tgacgtggcc tgcgatctca	2940
gtgacaatga tctgcttctg gatctcactg ttgcctttgg ttaggggaaca caactagtaa	3000
ctctgcagag tgccttctcc cgcagcccta ctggaacaca gcagagtctg tgccatgaag	3060
cagttacaga aacagaattg atgtgctgcc aaaaaaaaaa aaaaaatggg gcccgaata	3120
aaagaatata tagtactcac ctcagttcct tccataagaa gtgggtggtt taatgattgt	3180

taagccattt ttgcctgtgc cgggagcatg gagggctgag atgtcgacag gcagtgggaa 3240  
 acaaatgccc tcctaagcca caaggcgtgc gccagattag taggcaactc cattttaaga 3300  
 agctgccttt ttcacaaaac tggaagaaat aaaagcgggt ggaataaaca agttaaagt 3360  
 ctttaatgca aaaagtaatt gaaaggcagt gcctccattt tgggtgtactt tcttggaga 3420  
 aagtataaaa ttgaccggca tcatgagaga cggaagatgc cgtgttctca gccaaacaag 3480  
 caactctttc cccgccaggc actgtcgggt ggggtcaggc cagcttttaa acaactgggga 3540  
 ctggatcaca gaaaaacagt ggttttctgt ccctggaaat gaataggcac aaagaccac 3600  
 ttggctgtgg gcagactact cttcaataag atttgggtgg gaggaggaac attccttttg 3660  
 ctattttgag ctgagacaat ataaatatc aaactgtgcc atgcataaag cattgaattc 3720  
 tcagggcacc tcttcttccc cttaccctt ttaaggccat cccctccatt aataataatc 3780  
 caggtagttg tgaaaatcgt gcttctatct gatcccttct tagtttggtt tttcatccca 3840  
 tcagaacaag taaacgtagg cgccacagct cttgtgagta ctgtctccct cacggtgaat 3900  
 gagcctctg gtgtttcgtc caagaaaaga aagggtgtca ctggaaccac agcccttttt 3960  
 cattttataa actgcctctt catgttgcct gctcaagttt ccacctagaa ttgctatcac 4020  
 tgtggctctt tctaaaaatc tttctattta actggttcac tgaaattagt catagaaac 4080  
 ttgtgatttg gtgaagaggc attccttgta ataaccaa atgacttggg ggtgtgcata 4140  
 gcaagggcag tgttacactt atgaggactg tctctagcat ccaggaagtc tctgggtctg 4200  
 agggatggaa agttcttcct gctatgaatg agagtggact cttccctca cccccaactg 4260  
 aaaccacaaa caaccagaat cttctggaat tctgacttag agtcgttggt atagaagacc 4320  
 ttgttgctat ggaacatgaa actgtgtgtc agatggagag atcccttaa cctaagagcc 4380  
 ttaaatagcc ctgaaagtac actgggacgg tttgcatgg aattaaaatt ggaagtgaat 4440  
 attttttaggt gctcttgaag ctttctgggg actcaaaatt atcaaaagtc agggacagtc 4500  
 cggaggaaga gcgtctgcaa aactgggttc ctagaagtat agacggactt agctttttgt 4560  
 agaatttggg gaggagcagc gcctcgtgag agcagaatgg cctggcgtgg ccagtgttc 4620  
 ccgg 4624

<210> 116

<211> 3919

<212> DNA

<213> NM\_201524.1| Homo sapiens G protein-coupled receptor 56 (GPR56), transcript variant 2, mRNA

<400> 116

cgcactagct gtctgccctg ccctgccgta ggagatgggc tgggagcctc ccacgctctc 60  
 cagctcactc ggcaggcagc ggggaccagg gctggcaggt taagcctctg ggggtggatc 120

ctgaaagggtg	gtccagccgc	ctggccctgc	gtgggaccct	ccacctggca	gcagacaggg	180
tctcgctctg	tcacacaggg	tggagtgcag	tgggtgtgatc	ttggctcatc	gtaacctcca	240
cctccccgggt	tcaagtgatt	ctcatgcctc	agcctcccga	gtagctggga	ttacaggtgg	300
tgacttccaa	gagtgactcc	gtcggaggaa	aatgactccc	cagtcgctgc	tgacagcgac	360
actgttcctg	ctgagtctgc	tcttcctggg	ccaagggtgcc	cacggcaggg	gccacagggg	420
agactttcgc	ttctgcagcc	agcggaaacca	gacacacagg	agcagcctcc	actacaaacc	480
cacaccagac	ctgcgcatct	ccatcgagaa	ctccgaagag	gccctcacag	tccatgcccc	540
tttccttgca	gcccaccctg	cttcccgatc	cttccttgac	cccagggggc	tctaccactt	600
ctgcctctac	tggaaaccgac	atgctgggag	attacatctt	ctctatggca	agcgtgactt	660
cttgctgagt	gacaaagcct	ctagcctcct	ctgcttccag	caccaggagg	agagcctggc	720
tcagggcccc	ccgctgttag	ccacttctgt	cacctcctgg	tggagccctc	agaacatcag	780
cctgcccagt	gccgccagct	tcaccttctc	cttcacagt	cctccccaca	cggccgctca	840
caatgcctcg	gtggacatgt	gcgagctcaa	aagggacctc	cagctgctca	gccagttcct	900
gaagcatccc	cagaaggcct	caaggaggcc	ctcggctgcc	cccgccagcc	agcagttgca	960
gagcctggag	tcgaaactga	cctctgtgag	attcatgggg	gacatggtgt	ccttcgagga	1020
ggaccggatc	aacgccacgg	tgtggaagct	ccagcccaca	gccggcctcc	aggacctgca	1080
catccactcc	cggcaggagg	aggagcagag	cgagatcatg	gagtactcgg	tgctgctgcc	1140
tcgaacactc	ttccagagga	cgaaaggccg	gagcggggag	gctgagaaga	gactcctcct	1200
gggtggacttc	agcagccaag	ccctgttcca	ggacaagaat	tccagccaag	tcctgggtga	1260
gaagggtctg	gggattgtgg	tacagaacac	caaagtagcc	aacctcacgg	agcccgtggg	1320
gctcactttc	cagcaccagc	tacagccgaa	gaatgtgact	ctgcaatgtg	tgttctgggt	1380
tgaagacccc	acattgagca	gcccggggca	ttggagcagt	gctgggtgtg	agaccgtcag	1440
gagagaaacc	caaacatcct	gcttctgcaa	ccacttgacc	tactttgcag	tgctgatggg	1500
ctcctcgggtg	gagggtggacg	ccgtgcacaa	gcactacctg	agcctcctct	cctacgtggg	1560
ctgtgtcgtc	tctgccctgg	cctgccttgt	caccattgcc	gcctacctct	gctccaggag	1620
gaaacctcgg	gactacacca	tcaagggtgca	catgaacctg	ctgctggccg	tcttctgct	1680
ggacacgagc	ttcctgctca	gcgagccggg	ggccctgaca	ggctctgagg	ctggctgccg	1740
agccagtgcc	atcttccctg	acttctccct	gctcacctgc	ctttcctgga	tgggcctcga	1800
ggggtacaac	ctctaccgac	tcgtgggtgga	ggctcttggc	acctatgtcc	ctggctacct	1860
actcaagctg	agcgccatgg	gctggggcct	ccccatcttt	ctgggtgacgc	tggtggccct	1920
gggtggatgtg	gacaactatg	gccccatcat	cttggtctgtg	cataggactc	cagagggcgt	1980
catctaccct	tccatgtgct	ggatccggga	ctccctgggc	agctacatca	ccaacctggg	2040
cctcttcagc	ctggtgtttc	tgttcaacat	ggccatgcta	gccaccatgg	tggtgcagat	2100

cctgcggctg	cgccccca	cccaaaagt	gtcacatgtg	ctgacactgc	tgggcctcag	2160
cctgggtcctt	ggccttgcct	gggccttgat	cttcttctcc	tttgcttctg	gcaccttcca	2220
gcttgctcgtc	ctctaccttt	tcagcatcat	cacctccttc	caaggcttcc	tcctcttcat	2280
ctggtactgg	tccatgcggc	tgcaggcccc	gggtggcccc	tccccctctga	agagcaactc	2340
agacagcgcc	aggctcccca	tcagctcggg	cagcacctcg	tccagccgca	tctaggcctc	2400
cagcccacct	gcccattgtga	tgaagcagag	attcggcctc	gtcgcacact	gcctgtggcc	2460
cccagagccc	gcccagcccc	aggccagtca	gccgcagact	ttggaaagcc	caacgaccat	2520
ggagagatgg	gccgttgcca	tgggtggacgg	actcccgggc	tgggcttttg	aattggcctt	2580
ggggactact	cggctctcac	tcagctccca	cgggactcag	aagtgcgccg	ccatgctgcc	2640
tagggacttg	tccccacatc	tgtcccaacc	cagctggagg	cctgggtctct	ccttacaacc	2700
cctggggcca	gccctcattg	ctggggggcca	ggccttggtat	cttgagggtc	tggcacatcc	2760
ttaatcctgt	gcccctgcct	gggacagaaa	tgtggctcca	gttgctctgt	ctctcgtggg	2820
caccctgagg	gcactctgca	tcctctgtca	ttttaacctc	agggtggcacc	cagggcgaat	2880
ggggcccagg	gcagaccttc	agggccagag	ccctggcgga	ggagaggccc	tttgccagga	2940
gcacagcagc	agctcgccta	cctctgagcc	caggccccct	ccctccctca	gccccccagt	3000
cctccctcca	tcttccctgg	ggttctcttc	ctctcccagg	gcctccttgc	tccttcgttc	3060
acagctgggg	gtccccgatt	ccaatgctgt	tttttgggga	gtggtttcca	ggagctgcct	3120
ggtgtctgct	gtaaatgttt	gtctactgca	caagcctcgg	cctgccccctg	agccaggctc	3180
ggtaccgatg	cgtgggctgg	gctagggtccc	tctgtccatc	tgggcctttg	tatgagctgc	3240
attgcccctt	ctcaccctga	ccaagcacac	gcctcagagg	ggccctcagc	ctctcctgaa	3300
gccctcttgt	ggcaagaact	gtggaccatg	ccagtcccgt	ctggtttcca	tcccaccact	3360
ccaaggactg	agactgacct	cctctgggtga	cactggccta	gggcctgaca	ctctcctaag	3420
aggttctctc	caagccccc	aatagctcca	ggcgccctcg	gccgcccata	atggttaatt	3480
ctgtccaaca	aacacacacg	ggtagattgc	tggcctgttg	taggtggtag	ggacacagat	3540
gaccgacctg	gtcactcctc	ctgccaacat	tcagtctggg	atgtgaggcg	tgcgtgaagc	3600
aagaactcct	ggagctacag	ggacagggag	ccatcattcc	tgcctgggaa	tcctggaaga	3660
cttcctgcag	gagtcagcgt	tcaatcttga	ccttgaagat	gggaaggatg	ttctttttac	3720
gtaccaattc	ttttgtcttt	tgatattaaa	aagaagtaca	tgttcattgt	agagaatttg	3780
gaaactgtag	aagagaatca	agaagaaaaa	taaaaatcag	ctgttgtaat	cacctagcaa	3840
actggaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3900
aaaaaaaaaa	aaaaaaaaaa					3919

<210> 117

<211> 7401

<212> DNA

<213> NM\_014615.1| Homo sapiens KIAA0182 protein (KIAA0182), mRNA

<400> 117  
aagcagttta gacaaacact gggcgacggt ggctccagca tgtatcagcc gaggtggagc 60  
tgcgggggccc tggcatgaaa ggcatgagcc atgagcccaa gtcccccttcg ctaggggatgc 120  
tttcaccgcg gaccaggacc accgccaccg tcaaccccct cccccctcg ccgctcaatg 180  
gcgccctggt gcccagcggc agccccgcca ccagcagcgc gctgtcggcc caggccgcgc 240  
catectccag ctttgccgcc gcgctgcgca agctcgccaa acaggcggag gagcccagag 300  
ggtcctcact gagcagcggag tcgtcccccg tgctctctcc ggccaccaac cacagctccc 360  
ccgccagcac acccaagcgc gtgcccattg gccctatcat cgtccccctt gggggccaca 420  
gcgtgcccag ccccccccc gtggtgacca tcgtctcaac caaaaccgtg aatggtgtct 480  
ggaggagtga gagccggcag gatgccggct ccaggagcag cagtggaggt cgggaacgcc 540  
tcattgtgga gcccccgctc cctcaggaga aggcaggggg accagccatc ccctcgcacc 600  
tgctcagcac cccctacccc ttcggcctct cccccagctc agttgtgcag gattcccgcct 660  
tcccgccact caacctccag cgccccgtgc accacgtggt gccccccagt accgtgaccg 720  
aggactacct gagaagcttc cggccctacc acaccaccga cgacctccgc atgtcctcac 780  
tgctccccct cggcctggac ccggccactg ctgcagccta ctaccacccc agctacctgg 840  
ccccacaccc cttccccac ccggccttca ggatggacga ctctactgc ctgtctgccc 900  
tgagggtccc gttctacccc atccccaccc ccggctccct gccccactg caccatcag 960  
cgatgcacct gcacctctct ggggtccgct accctcccga gctctccac tcatccctgg 1020  
cagcgctgca ctcgagcgc atgtctggcc tcagcgcgga gaggctgcag atggacgagg 1080  
agctaaggcg ggagagggag cgcgagcgc agcgcgagcg tgagcgtgag gctgaccgcg 1140  
agcgggagaa ggaacgtgag cgcaacgcg agaaggagcg cgagcaagag aaggagcgtg 1200  
agcgtgagaa ggagcgcgag cgcgagctgg agcgcagcg ggagcagcg gcccgggaga 1260  
aggagctgct ggccgccaag gccctggagc ccagcttcct gcccgtggcc gagctgcatg 1320  
ggctgcgtgg ccatgccact gaggagcggg gcaagccctc ggagcagctg accccaaccc 1380  
gagcagagaa gctgaaggat gccggcctgc aggcgcccga gcccgtccaa cacccttgc 1440  
atccggtgcc cccccacac cacacggtgc ccagcctcat ctccaacat ggcattctct 1500  
ctctgcctag cagcagtgtc gccacagccc tgctgatcca gcgcaccaat gaggaggaga 1560  
agtggctggc ggggcagcgg cggctgcggc aggagaagga ggaccggcag tctcaggtgt 1620  
ccgagttccg gcagcaggtg ctggagcagc acctggatat gggccggccc ccggtgccgg 1680  
cggaggcaga gcacaggccg gagagcacca ccaggccagg accaaaccgt cacgagccag 1740  
gtggccgtga ccctccgcag cactttgggg ggccaccacc tctgatttcg cccaagcccc 1800

agctccatgc	tgcacccacg	gcccctctgga	accccgtgtc	cctgatggac	aacaccttgg	1860
agacgcggcg	ggccgaaagc	cactctctgc	acagccaccc	ggctgcattt	gagcccagcc	1920
gccaggcagc	cgtgccgctg	gtgaaggtgg	agcgggtctt	ctgcccggag	aaagcagagg	1980
agggggccacg	gaagcgtgag	cctgccccctc	tggacaagta	ccagccacct	ccgccgccac	2040
cacgagaggg	agggagcctg	gagcaccagc	ccttcctgcc	cgggcccggg	cccttcctgg	2100
ctgagctcga	gaagtccacc	cagaccatcc	tgggccagca	gcgggcctcc	ctcccacagg	2160
cggccacctt	cggggagctc	agcggacccc	tgaagcctgg	ctcgccctac	cggccccag	2220
tgccacgggc	ccccgacctt	gcctacatct	atgatgagtt	cctgcagcag	cgccggaggc	2280
tggtcagcaa	gctggacctg	gaggagcgca	ggcggcgggg	ggcccaggag	aaaggtact	2340
actacgacct	cgatgactct	tacgacgaga	gcgatgagga	ggaggtcagg	gcccacctcc	2400
gttgctggc	cgagcagccg	cccctcaaac	tggacacgtc	ctctgagaag	ctagagtttt	2460
tgcaactttt	tggcttgacc	acccaacagc	agaaggagga	attggtggcc	cagaagcgga	2520
ggaagcggcg	gaggatgctg	cgagagagaa	gcccgtcgcc	cccaacaatt	cagagcaagc	2580
ggcagacgcc	ttcaccgaga	ctggcgctgt	ctaccgccta	cagccctgat	gagatgaaca	2640
acagtcccaa	cttcgaagaa	aagaagaagt	tcctgacct	cttcaacctg	accacatca	2700
gcgctgagaa	gaggaaaagac	aaagagagac	ttgttgaaat	gctccgtgcc	atgaagcaga	2760
aggcactgtc	agcagcagtg	gccgactcct	tgacaaactc	tccgagggac	agtcctgccg	2820
tctccctgag	tgaaccagcc	acgcagcaag	cctctctgga	tgtggagaag	ccggttggtg	2880
ttgtgcttc	cttgtctgac	atcccaaagg	ccgcggagcc	tgggaagctg	gaacaggtcc	2940
ggccccagga	gctgtcgaga	gtccaggagc	tagctcctgc	cagcggggag	aaggccaggc	3000
tgagcgaggc	ccctggaggc	aaaaagagtc	tgagcatgct	tcactatata	cggggcgctg	3060
cacccaagga	cattcctgtg	ccgctgtccc	acagcaccaa	tgggaagagc	aagccgtggg	3120
agccctttgt	ggcagaagag	tttgacatac	agttccacga	gtcagtgtctg	cagtccaccc	3180
agaaggccct	gcagaagcat	aaagggagcg	tggctgtgct	gtctgcagag	cagaaccaca	3240
aggttgacac	gtccgtccac	tacaacattc	ctgagctgca	gtcctccagc	cgcgccccctc	3300
caccccagca	caatgggcag	caggagcccc	ccactgcaag	gaagggcccc	ccaaccagg	3360
agttggaccg	ggactcggag	gaggaggaag	aggaggatga	tgaagatgga	gaagatgagg	3420
aggaagtccc	caagcgcaag	tggcaaggga	tcgaggccgt	ttttgaagct	taccaggaac	3480
acatagaaga	gcaaaatctg	gagcggcagg	tgttacagac	acaatgtaga	cgactggagg	3540
cccggcacta	cagcctcagc	ctgacggcag	agcagctctc	ccacagcgtg	gcggagttga	3600
ggagccagaa	acagaagatg	gtctcagaaa	gggagcggct	ccaggcagaa	ctggaccact	3660
tacgaaagtg	ccttgccctg	cctgcaatgc	actggcctag	gggctacctg	aagggatata	3720
ccaggtgacg	gtttcccttg	cactaggccg	aacctatagt	atagaaatat	tatctatatt	3780
attaccttga	atatttaata	tttttactg	ggaggtttga	agcttacaaa	atgagaatgt	3840



gccatgcatg	aagcaaagga	ttccaggctc	cagaaaaaat	gaatgaactc	accttgacgt	3900
caatgcaatt	gaatcacccgt	tgtcattcag	cgagcaacca	atgtaggatt	gccacagtt	3960
tttcttttta	aagggtggtt	tgcaccttcc	tctcccat	tatttcttaa	tctgaacatg	4020
aaggctccat	tagcaacact	aaaacttgat	cattaacagc	cccctgtgca	tatgagtgg	4080
tcaaaccggt	tctgttcttt	cttgtgttgc	catgttacta	tgcctcaagc	ccagtttgct	4140
tttgccgcag	cgatggggcc	agtctcattc	ctccccagga	gtgaaacttg	cttcagctga	4200
aaagggtggg	tgcatgtgca	gtaaaaaggg	cttatttggt	tcattttact	ttcctgcaaa	4260
attttcttca	aagcaacaag	tcctaggagc	acacaaagca	acccaaaggc	ttttccctgg	4320
aaaagctctt	tcttacctaa	agataaaacc	aattcacaaa	ctgaaggtag	ctttttatta	4380
ctccgtgggg	agcatgtaca	gagctctgtg	tatacacagc	ttcacacca	ccagattggt	4440
actacagtgg	gttgggtttt	catacagacg	taaattttga	gagaaaagtc	aaagggtgctt	4500
cagccttgta	ctgtgtatat	atattaaaaa	aaaaacaaag	ttttgtatgt	ttttattact	4560
ttaactattg	ttataaaaag	cctgccattt	ttaatattgt	gtttggggga	ttttgtttg	4620
tttttcctgt	ttgggggttt	tgtttggtgt	tttggtttt	tttgggcaaa	aaaaaaaaaa	4680
aaaccttgct	tttagtgttt	gtactgctgc	tggtcaggac	attaaaatat	tgaagtgttt	4740
ttaaaaatta	aagaagaaga	aaagtaaaag	agcttaccac	tggcgccctat	gcatcactt	4800
catttttagt	ttgagttgca	ccagaagctg	ccgtagaaag	ccatgcgcta	ctgcttacct	4860
cctccactcc	ccctgcctgc	ccccagcatc	tggacaagct	aatagcaa	attaccatt	4920
gctatcaagg	gaggaggggg	tagtctgtag	aacccatgtg	tgacagtc	gtgcacacat	4980
gggcgggggc	ttttaaaaac	ctttcaggaa	gtcaatgatt	tctgtgattg	atataattct	5040
aagggtgtctg	agagcaggta	cagaatagga	acttcagagg	ctttgtttta	acgcaaagct	5100
ttgtaaaagc	cacaaggtct	gagctgaacc	cctccttttt	gaacttactg	tgacaagcac	5160
aggaacggtc	agaaactggg	ctcatcacac	caaggcaaag	caacgggcga	gtcttcctcc	5220
ttgtcctagt	tactgcctat	ggaggcagtg	tttagatcaa	gaaggcctct	cttgctccca	5280
agggccctca	ccagaggcca	gggctgccag	tactggtct	ggggggtgga	ggcctgagct	5340
gagggcaggg	tgctgacct	gtgtgccggc	tgctcactgc	tgtgaccagc	agccgagccc	5400
ttggccctag	cccttgctgc	gcagaacagc	ttgctggcag	ctggcatcgt	gtcgctttat	5460
ctgccccgc	acagtttgct	ttgtacgtct	gccagaatc	ttccagttat	tagcaaactc	5520
agacgaatgt	accgccagta	ttatcagcag	tcaacaagca	ccttcctctc	cacagaagca	5580
gctggaagag	aactcgaggg	gctgtgctgc	aggcctcccc	tcgaaagaca	ctgggaggctc	5640
agcatgttcc	acagggtgtt	agagggagtc	tgctacaaac	tatcagggca	aaatctcact	5700
ggatttctcc	actgaaaacc	tacttgaggt	ttctgggtctg	aaggcttaag	agtcacatct	5760
tagcacttcc	gctctcaggc	ctcctcctcc	atcacagatg	tctggatgct	tttggaaatg	5820

gccttggtcta	aagtaaaagg	gaaaagtaga	tccgataact	taaaaacgta	gctcatccct	5880
taccatccaa	ggggcactcc	cttggttga	ttttctatga	cagcacaggg	gacaggtggc	5940
acaccatgag	aggtctgccc	aggggtgggag	cagtgtcact	gtgctagcaa	tagttggctt	6000
ctccccgtc	agtggaaacc	ccacttctgc	ccggcccttg	agcttcttgc	ccactgtctc	6060
cccattcttc	cacctacttg	tggcgatctg	agtactctac	tcttgctcaa	gaagtaatac	6120
gacaatcaga	atacaaacca	gtaaggcaac	acgaataaac	taagaaaaag	gtaagaactg	6180
tctcaaaaac	gaaagcacac	cacccaagac	acagtaccca	gtcatggttt	ccccatccaa	6240
ctattagttt	catactttga	aaacttactt	tcagattatt	ctcaaagaac	acagtagcac	6300
ctaaatctgt	tttcaattgg	gcttaaaaat	tgacatgcaa	tctcttaagt	tttttgttca	6360
gctacttcac	actgagtacc	tcaaactctgc	tctggagtcg	attatgccac	ctgtgtgtca	6420
ggatgcacct	gaaagccctc	ggctcgggtcc	ttagaccatc	ttcctacatt	acctggaagg	6480
gagctgccat	ctgtccctct	gcagagggat	accttccaat	agtaaattat	ctggttcctc	6540
actgaaacaa	gttatttttg	cttcatatag	tcagagtcag	actgacatga	taaaatatca	6600
tgttccta	ctgttgcttc	agataagtga	ccaagacggg	actttccaca	ttttagtcta	6660
cattctaatt	ttaaaggaat	aaagcactga	attgggacta	acattctgat	aggttgcacc	6720
cttaagagta	ttcagagagc	atcaaaagga	gccacacct	tcagcagtga	aggattctaa	6780
cacaggggaat	ctgcagtttg	tagcagaatg	gtattttcct	caagtagctc	ataatactgc	6840
caaatctcaa	aagttaagct	gaatttcaca	ccagatccta	cccctttccc	tgagccacat	6900
gtttcacaca	agtgtagaaa	atgccagggg	tccaccacaa	gatggagatg	gtcagcacaa	6960
accgattctg	ttcctcttta	aagtgtatat	tagccactta	gcaatctcta	tattctttca	7020
agtaaccaag	ctgttgactt	tcttactact	tgcagtagcc	tgtccccaac	ttttccatcc	7080
agtgtttaac	ctaaaaaact	ccttaactct	gccttgacct	gaggaagacc	atgctaactg	7140
gtgttatatt	gtatgtaccc	tgtgtttaat	tctataacag	taaaccccat	acgcaggtgg	7200
gagggaggaa	caccgggtgcc	tcggtcactc	tgggggcagt	ttagatgctg	tgaaattaaa	7260
cctgttctaa	gtgtacttgt	ttgaattaat	tgtattgtaa	tattatttgt	tgaatgtagt	7320
aattaggtat	ttatgaatat	attgctgtaa	tttctgacaa	catccaaaaa	ataaaatctt	7380
cctaaattat	gttaaaaaaa	a				7401

<210> 118

<211> 2745

<212> DNA

<213> NM\_033542.1| Homo sapiens chromosome 20 open reading frame 35 (C20orf35), mRNA

<400> 118

tttttccccg gaaacgtttc tttcctacgc agccgctcct gccgccgtgg tcgctggagc	60
tttgctctc taggccggca ggcctctcc tccatgggcc tgtctgtcag cgctgttttg	120
ggagcccgcc ggtgaggccg ggccacgctc agacacttcg atcgctcagc ctgtcactgg	180
gcatggcggg tcagttccgc agctacgtgt gggacccgct gctgatcctg tcgcagatcg	240
tcctcatgca gaccgtgtat tacggctcgc tgggctgtg gctggcgctg gtggacgggc	300
tagtgcaag cagccccctc ctggaccaga tgtttgacgc cgagatcctg ggcttttcca	360
ccccccagg ccggctctcc atgatgtcct tcctcctcaa cgcctcacc tgtgccctgg	420
gcttgctgta ctcatccgg cgaggaaagc agtgtctgga tttcactgtc actgtccatt	480
tctttcacct cctgggctgc tgggtctaca gctcccgttt cccctcggcg ctgacctggt	540
ggctgggtcca agccgtgtgc attgcactca tggctgtcat cggggagtac ctgtgcatgc	600
ggacggagct caaggagata cccctcaact cagcccctaa atccaatgtc tagaatcagg	660
ccctttggac atcctgctga cacttggggc ccttaacacc ttgggctgct cagaccctcc	720
agatgagggtc cagcccagat ctgagaggaa ccctggaaat gtgaagtctc tgttggtttg	780
ggagagatag tgagggcctg tcaaagaagg caggtagcag tcagcatgac agctgcaaga	840
atgacctctg tctgttgaag ccttgggtatc tgagagggtca ggaaggggac ctctttgagg	900
gtaataacag aattggaacc atgccactct tgagccacaa tacctgtcac cagcctgttg	960
ttttaagaga gaaaaaaaaat caaggatatc tgattggagc aaaccacttc tttagtcatc	1020
tgtcttacc ccctgggaca gctgttacct ttgcagtgtt gccgaatcac agcagttacc	1080
tttgagctgt tgccgaatca cagcagttct gttggagaaa cgcttggttt ccggatccag	1140
agccacagaa agaaatgtag gtgtgaagta ttaggctgct gtcaggggaga ggatggcaga	1200
tggaggcatc aagcacaagg aaaatgcaca acctgtgccc tgttatacac acgttcatgt	1260
gcaccaaga acctatgact ttcttccagt tccttctacc aggtcccat cctgctgcca	1320
gctctcaaca tagcaggcca taggaccag agaagaatcc cagcgttgct caaagtctaa	1380
ccatcataaa gacactgcct gtcttctagg aatgaccagg caccagctc ccactggact	1440
ccaatttttt ttctgcctt atttagaatt ctttgccggg aagggtatga tgggttccca	1500
gagacaagaa gcccaacctt ctggcctggg ctgtgctgat agtgctgagg gagataggaa	1560
tttgctgcta agatttttct ttgggggtgga gtttctctg tgaggggctt gcagctatcc	1620
ttcctgtgta taaaaataca gtattttcca tgggtctgcc tgcacttact ttgtaatgcc	1680
acggttgaga ttgagagaga tcagcgcagc caggcaaggg aactttaaag aattattagg	1740
ccaccttctc ctttctctgg accccagagt cattcctcca tttggttaaa atactcagtg	1800
cagggaactc ttacatcctg tctccttcac ttgcagcgtc ccctgctatg cctcagggtga	1860
accacataat tcttgggttt ccgttcttac ttgctagtga tttctgaaca tgttcaatgg	1920
agcggcacac agtctagacc cacttccgca ttgaaacctt cactgttctt ctttgggttc	1980
ttcagagctt tcccaagaga gctgtcagtt ttcagctgtc agtaacacaa atgagtttat	2040

ggtaacacaa atgagttttg ctatctctct gagaagctca tctgacctcc tgactctcag	2100
ccctacagag tagggagttg atgctgacag gatgaagatt taggaataaa tatgcctggg	2160
aagagactgg gaaggttcta gggtgaggca cctcagtaac tcatggtacc ttggccaagt	2220
tggaaggaag cagtttgttt atgaggcaca gtaatcctgg ctgcagggtc taggaggtaa	2280
gaccagctgg gatgaccttc cctgggttaa tcaattttccc tctagacaac acaaactgca	2340
ggcatgtgac taactttgaa agaacaccca tcatgtggct gctgtcacc ttgaccagcc	2400
gtggtggtgg ttactccatc tgtggttgga gcgcctcttt gggattcact tcaaggctct	2460
gtgcctatct ttctgcatat cttctgtgat gacaaatctc tgtcccctga gtgttaatct	2520
gattttttaga aatggccaaa agtcacgtga tccaaacttt ttttcagtaa tatggagact	2580
gagctgcatg gtagttgggg atcaaaaata tgtgacctta atgagatttt tatgatttct	2640
aaagtaacaa taaaagcagc ttttagagtt gagttccaga gagggcaggg caatggcagt	2700
gacatgtttg tcattttaat aataaataac atctattgag tgctt	2745

<210> 119

<211> 2152

<212> DNA

<213> NM\_138932.1| Homo sapiens apobec-1 complementation factor (ACF), transcript variant 2, mRNA

<400> 119	
tttgatatga cgattagagc ataacccgag tgacacgttg aattcgccat aatcaaggaa	60
accttttccg ggtggggatc tctgaaatta ctcagataac agtgctgtgc caaaaacctg	120
tggattttct ctacaaaaat tattgagcaa ccctaattaa cctgattttt tgctgataat	180
cactctcaat ggaatcaat cacaaatccg gggatggatt gagcggcact cagaaggaag	240
cagccctccg cgcactggtc cagcgcacag gatatagctt ggtccaggaa aatggacaaa	300
gaaaatatgg tggccctcca cctggttggg atgctgcacc ccctgaaagg ggctgtgaaa	360
tttttattgg aaaacttccc cgagaccttt ttgaggatga gcttatacca ttatgtgaaa	420
aaatcggtaa aatttatgaa atgagaatga tgatggattt taatggcaac aatagaggat	480
atgcatttgt aacattttca aataaagtgg aagccaagaa tgcaatcaag caacttaata	540
attatgaaat tagaaatggg cgcctcttag gggtttgtgc cagtgtggac aactgccgat	600
tatttgttgg gggcatccca aaaacaaaa agagagaaga aatcttatcg gagatgaaaa	660
aggttactga aggtgttgtc gatgtcatcg tctaccaag cgctgcagat aaaacaaaaa	720
accgaggctt tgccttcgtg gagtatgaga gtcatcgagc agctgccatg gcgaggagga	780
aactgctacc aggaagaatt cagttatggg gacatggtat tgcagtagac tgggcagagc	840
cagaagtaga agttgatgaa gatacaatgt cttcagtga aatcctatat gtaagaaatc	900

ttatgctgtc	tacctctgaa	gagatgattg	aaaaggaatt	caacaatatc	aaaccaggtg	960
ctgtggagag	ggtgaagaaa	attcgagact	atgcttttgt	gcacttcagt	aaccgagaag	1020
atgcagttga	ggctatgaaa	gctttaaatg	gcaaggtgct	ggatggttcc	cccattgaag	1080
tcaccctagc	aaaaccagtg	gacaaggaca	gttatgttag	gtatacccga	ggcacaggtg	1140
gaaggggcac	catgctgcaa	ggagagtata	cctactcttt	gggccaagtt	tatgatccca	1200
ccacaacctt	ccttggagct	cctgtcttct	atgcccccca	gacctatgca	gcaattccca	1260
gtcttcattt	cccagccacc	aaaggacatc	tcagcaacag	agccattatc	cgagcccctt	1320
ctgttagaga	aattttacatg	aatgtacctg	taggggctgc	gggagtgaga	ggactgggctg	1380
gccgtggcta	tttggcatac	acaggcctgg	gtcgaggata	ccaggtcaaa	ggagacaaaa	1440
gagaagacaa	actctatgac	attttacctg	ggatggagct	cacccaatg	aatcctgtca	1500
cattaaaacc	ccaaggaatt	aaactcgctc	cccagatatt	agaagagatt	tgtcagaaaa	1560
ataactgggg	acagccagtg	taccagctgc	actctgctat	tggacaagac	caaagacagc	1620
tattcttgta	caaaataact	attcctgctc	tagccagcca	gaatcctgca	atccaccctt	1680
tcacacctcc	aaagctgagt	gcctttgtgg	atgaagcaaa	gacgtatgca	gccgaataca	1740
ccctgcagac	cctgggcatac	cccactgatg	gaggcgatgg	caccatggct	actgctgctg	1800
ctgctgctac	tgttttccca	ggatatgctg	tccctaatagc	aactgcaccc	gtgtctgcag	1860
cccagctcaa	gcaagcggta	acccttggac	aagacttagc	agcatataca	acctatgagg	1920
tctacccaac	ttttgcagtg	actgcccag	gggatggata	tggcaccttc	tgaagatgct	1980
tttttaaatt	taagaataag	acacacaaaa	ctctattaaa	aaaaaaaaag	aaataaacct	2040
ctaactcgg	ccccaatgat	cataaataat	atgtttccta	aagaaatgcc	tttccagaga	2100
ctgtatagct	tataccaatt	atagaatcat	gaagtaaaaa	aaaaaaaaaa	aa	2152

<210> 120

<211> 3010

<212> DNA

<213> NM\_145343.1| Homo sapiens apolipoprotein L, 1 (APOL1), transcript variant 2, mRNA

<400> 120	
actttccctt	tcgaattcct cggtatatct tggggactgg aggacctgtc tggttattat 60
acagacgcat	aactggaggt gggatccaca cagctcagaa cagctggatc ttgctcagtc 120
tctgccaggg	gaagattcct tgacttctgg ggtgatggag aagaaacagg ctgtgctgtg 180
tccctaattg	gaaacgtggc tgagacaggg gagtgagaag ggtgcgttgc agaattggtgc 240
ctgtggcatg	atgccagctt tgcaatcatg agattcaaaa gccacactgt ggaattgagg 300
agggccctgc	gcgacatgga gggagctgct ttgctgagag tctctgtcct ctgcatctgg 360

atgagtgcac	ttttccttgg	tgtgggagtg	agggcagagg	aagctggagc	gaggggtgcaa	420
caaaacgttc	caagtgggac	agatactgga	gacacctcaa	gtaagcccct	cggtgactgg	480
gctgctggca	ccatggaccc	agagagcagt	atctttattg	aggatgccat	taagtatttc	540
aaggaaaaag	tgagcacaca	gaatctgcta	ctcctgctga	ctgataatga	ggcctggaac	600
ggattcgtgg	ctgctgctga	actgcccagg	aatgaggcag	atgagctccg	taaagctctg	660
gacaaccttg	caagacaaat	gatcatgaaa	gacaaaaact	ggcacgataa	aggccagcag	720
tacagaaaact	ggtttctgaa	agagtttcct	cggttgaaaa	gtgagcttga	ggataacata	780
agaaggctcc	gtgcccttgc	agatgggggt	cagaaggctc	acaaaggcac	caccatcgcc	840
aatgtgggtg	ctggctctct	cagcatttcc	tctggcatcc	tgaccctcgt	cgcatgggt	900
ctggcacctt	tcacagaggg	aggcagcctt	gtactcttgg	aacctgggat	ggagttggga	960
atcacagccg	ctttgaccgg	gattaccagc	agtaccatgg	actacggaaa	gaagtgggtg	1020
acacaagccc	aagcccacga	cctgggtcatc	aaaagccttg	acaaattgaa	ggaggtgagg	1080
gagtttttgg	gtgagaacat	atccaacttt	ctttccttag	ctggcaatac	ttaccaactc	1140
acacgaggca	ttgggaagga	catccgtgcc	ctcagacgag	ccagagccaa	tcttcagtca	1200
gtaccgcatg	cctcagcctc	acgcccccg	gtcactgagc	caatctcagc	tgaaagcgg	1260
gaacaggtgg	agaggggtta	tgaacccagc	atcctggaaa	tgagcagagg	agtcaagctc	1320
acggatgtgg	cccctgtaag	cttcttttct	gtgctggatg	tagtctacct	cgtgtacgaa	1380
tcaaagcact	tacatgaggg	ggcaaagtca	gagacagctg	aggagctgaa	gaaggtggct	1440
caggagctgg	aggagaagct	aaacattctc	aacaataatt	ataagattct	gcaggcggac	1500
caagaactgt	gaccacaggg	cagggcagcc	accaggagag	atatgcctgg	caggggccag	1560
gacaaaatgc	aaactttttt	ttttttctga	gacagagtct	tgctctgtcg	ccaagttgga	1620
gtgcaatgg	gcgatctcag	ctcactgcaa	gctctgcctc	ccgtgttcaa	gcgattctcc	1680
tgccttgggc	tccaagtag	ctgggactac	aggcgcctac	caccatgccc	agctaatttt	1740
tgtattttta	atagagatgg	ggtttcacca	tgttggccag	gatggctctg	atctcctgac	1800
ctcttgatct	gcccaccttg	gcctcccaaa	gtgctgggat	tacaggcgtg	agccatcgct	1860
tttgacccaa	atgcaaacat	tttattaggg	ggataaagag	ggtgaggtaa	agtttatgga	1920
actgagtgtt	agggactttg	gcattttccat	agctgagcac	agcaggggag	gggttaatgc	1980
agatggcagt	gcagcaagga	gaaggcagga	acattggagc	ctgcaataag	ggaaaaatgg	2040
gaactggaga	gtgtggggaa	tgggaagaag	cagtttactt	tagactaaag	aatatatattg	2100
ggggccgggt	gtagtggctc	atgcctgtaa	tccgagcact	ttgggaggcc	aaggcggg	2160
gatcacgagg	tcaggagatc	gagaccatcc	tggctaacac	agtgaaaccc	cgtctctact	2220
aaaaatacaa	aaaattagcc	gggcatgggt	gcgggcgcct	gtagtccag	ctaactgggc	2280
ggctgaggca	ggagaatggc	gtgaacctgg	gaggtggagc	ttgcagtgag	ccgagatatc	2340

gccactgcac	tccagcctgg	gtgacagagc	gagactccat	ctcaaaaaaa	aaaaaaaaaa	2400
gaatatattg	acggaagaat	agagaggagg	cttgaaggaa	ccagcaatga	gaaggccagg	2460
aaaagaaaga	gctgaaaatg	gagaaagccc	aagagttaga	acagttggat	acaggagaag	2520
aaacagcggc	tccactacag	acccagcccc	aggttcaatg	tcctccgaag	aatgaagtct	2580
ttccctgggtg	atgggtcccc	gccctgtctt	tccagcatcc	actctccctt	gtcctcctgg	2640
gggcatatct	cagtcaggca	gcggttcctt	gatgatggtc	attgggggtg	ttgtcatgtg	2700
atgggtcccc	tccaggttac	taaagggtgc	atgtcccctg	cttgaacact	gaagggcagg	2760
tgggtgggcca	tggccatggt	ccccagctga	ggagcagggtg	tccttgagaa	cccaaacttc	2820
ccagagagta	tgtgagaacc	aaccaatgaa	aacagtccca	tcgctcttac	ccggtaagta	2880
aacagtcaga	aaattagcat	gaaagcagtt	tagcattggg	aggaagctca	gatctctaga	2940
gctgtcttgt	cgccgcccag	gattgacctg	tgtgtaagtc	ccaataaact	cacctactca	3000
tcaagctgga						3010

<210> 121

<211> 2759

<212> DNA

<213> NM\_080796.1| Homo sapiens death associated transcription factor 1 (DATF1), transcript variant 2, mRNA

<400> 121						
tctagccccg	cgccatctcg	gtggccgtcc	gccactccg	cggcgttcgg	ggaaatggct	60
gcgagaccct	agaggcctgc	ggcctgcgga	gcttactcca	cggaacagc	ctctagataa	120
tctgagttgt	tgaaaatacg	aagcctgtta	ctcgtgaaca	gtggctgaca	acagtgttgt	180
tgtgagcctg	gctgtctgct	tggaccacaga	ggtttcgtct	gccagggttt	ttggttgtat	240
ttaggatttc	agggaaaagt	gtccaagctt	tcagtgttgg	agcaggatatg	gacgacaaag	300
gcgacccgag	caatgaggag	gcacctaagg	ccatcaaacc	caccagcaaa	gagttcagga	360
aaacatgggg	ttttcgaagg	accactatcg	ccaagcgaga	gggcgcaggg	gacgcggagg	420
ctgacccact	ggagccgcca	ccccacagc	agcagctggg	cctgtccctg	cggcgcagtg	480
ggaggcagcc	caagcgcact	gagcgcgtgg	agcagttcct	gaccattgctg	cggcgcgcgcg	540
gcaggaggag	catgcctgtc	tccctggagg	attctgggtga	gccacgtcc	tgccccgcca	600
cagacgccga	gacagcctcc	gagggcagcg	tggaaagcgc	ttctgagacc	agaagcggcc	660
cccagtctgc	ttccacagct	gtgaagggaac	gaccagcctc	ttctgaaaag	gtgaaaggag	720
gggatgacca	cgatgacacc	tccgatagtg	acagcgatgg	cctgaccttg	aaagagcttc	780
agaatcgctt	tcgcaggaag	cgggaacagg	agccactga	gaggccccctg	aaagggatcc	840
agagtcgcct	gcggaagaag	cgccgggagg	aggggtccgc	cgagactgtg	ggctccgagg	900

ccagtgcac	tgtggagggc	gtcctgcca	gtaagcagga	gcccgagaac	gatcagggg	960
tttgttcca	ggctgggaaa	gatgacagag	agagtaagtt	ggagggaaa	gcggctcagg	1020
acatcaaaga	tgaggagcct	ggagacttgg	gccgaccgaa	gcctgaatgt	gagggttacg	1080
acccaacgc	cctgtattgc	atttgcggcc	agcctcaca	caacagggtt	atgatttgct	1140
gtgaccgctg	tgaagaatgg	tttcatggcg	attgtgtggg	catttctgag	gctcgaggga	1200
ggcttttga	aaggaatggg	gaagactata	tctgccccaa	ctgcaccatt	ctgcaagtgc	1260
aggatgagac	tcattcagaa	acggcagatc	agcaggaagc	taaatggaga	cctggagatg	1320
ctgatggcac	cgattgtaca	agtataggaa	caatagagca	gaagtctagc	gaagaccaag	1380
ggataaagg	tagaattgag	aaagctgcaa	atccaagtgg	caagaagaaa	ctcaagatct	1440
tccagcctgt	gatagaggcg	cctgggtgcct	caaaatgtat	tggccccggg	tgctgtcacg	1500
tggcgcagcc	cgactcgggtg	tactgcagta	atgactgtat	cctcaaacac	gccgcagcga	1560
caatgaagtt	tctaagctca	ggtaaagaac	agaagccaaa	gcctaaagaa	aagatgaaga	1620
tgaagccaga	gaagcccagt	cttccgaaat	gcggtgctca	ggcaggtatt	aaaatctctt	1680
ctgtgcacaa	gagaccagct	ccagaaaaaa	aagagaccac	agtgaagaag	gcagtgggtg	1740
tccctgcgcg	gagtgaagca	ctcgggaagg	aagcagcttg	tgagagcagc	acgccgtcgt	1800
gggcgagcga	tcacaattac	aatgcagtaa	agccagaaaa	gactgctgct	ccctcgcctg	1860
cactgttgta	taaatgtatg	tatcacctag	gggttggcct	cctggacccc	tcccgttctt	1920
tctggatagc	catcccctgg	gcctgtccag	gactgggagt	tgagctttg	tggttaagctg	1980
atcacagaca	ccggctgcac	catcagcggg	aagcagagcc	catgtccagg	atgcctcctg	2040
ctgccctgtg	tccatcccta	gtctgtcagg	acttcctgtc	actgttttcc	aaagctgtaa	2100
acctcactgg	tgaacgttca	ccttaatgat	tgattcttta	atctctgttt	tcactctcag	2160
gctctggtaa	gtattcgtat	tctcttcac	ccagtctgat	tgcatagcca	cactgcccgg	2220
cacgccacat	ccaccctgt	ctgcacatga	gttgttctga	caacagcgct	gtatacgctt	2280
cagtttttcc	acattgtcca	cggccagcac	atgaaagcat	cacttctttt	ttatgtttgtg	2340
ggaatctttg	caagtttagtg	ttgcatctga	ttttcaggtg	tacatttatt	tttgactggg	2400
cagatagggg	atTTTTTTTT	ttccatgtcc	gattcacacg	ctacacaccc	acatgaacac	2460
attcgaactt	cgaaggccac	acactcctgc	ttcataggcc	ccacggtaag	tgagttcaca	2520
cctagaacac	tgctcctgacc	gcaggacgcg	tgccctggac	ttgggtattct	acatgtgact	2580
ggcttttctg	ccctcgtctc	ttgaatgttt	agactcttaa	gatcatatcc	tgccccaat	2640
ttcaaatata	tgaaatgaag	atatttcaaa	cagatctttg	aaacctcaga	ttctgtgggtg	2700
caattttaat	gttttcttgt	ttctcagttt	tctgctataa	aactattttc	aattcagtc	2759

<210> 122

<211> 781



<212> DNA

<213> NM\_177953.1| Homo sapiens dynein, cytoplasmic, light polypeptide 2A (DNCL2A), transcript variant 2, mRNA

<400> 122  
cgcagaaagg cacaggactc gctaagtgtt cgctacgcgg ggctaccgga tcggtcggaa 60  
atggctgaag tggagagatc gcctgagccc aggaggtcaa ggctacagtg agccgtgact 120  
gcaccactgc actccaccct gggcagaggt ggaggagaca ctgaagcgac tgcagagcca 180  
gaagggagtg caggggaatca tcgtcgtgaa cacagaaggc attcccatca agagcaccat 240  
ggacaacccc accaccaccc agtatgccag cctcatgcac agcttcaccc tgaaggcacg 300  
gagcaccgtg cgtgacatcg acccccagaa cgatctcacc ttccttcgaa ttcgctccaa 360  
gaaaaatgaa attatggttg caccagataa agactatttc ctgattgtga ttcagaatcc 420  
aaccgaataa gccactctct tggctccctg tgtcattcct taatttaatg ccccccaaga 480  
atgttaatgt caatcatgtc agtggactag cacatggcag tcgcttgga cccactcaca 540  
ccaatccagt gaccgtgtgt gggctggcgg ctcttctccc ccaccaacgg aaccctgtg 600  
tgcaccaacc ttccccagag ctccggagcg ccctctcctc acttccaggt tttggagcaa 660  
gagcttgtag gaagcccgcg ccagcttcc ttctgacctt cagttcactt tgtcgccctt 720  
ggagaaaagct gtttttcttt aactaaaaat aaccaaaatg cttaaaaaaa aaaaaaaaaa 780  
a 781

<210> 123

<211> 841

<212> DNA

<213> NM\_022873.1| Homo sapiens interferon, alpha-inducible protein (clone IFI-6-16) (G1P3), transcript variant 3, mRNA

<400> 123  
gaaccgttta ctgctgctg tgcccatcta tcagcaggct ccgggctgaa gattgcttct 60  
cttctctcct ccaagggtcta gtgacggagc ccgcgcgcgg cgccaccatg cggcagaagg 120  
cggatcgct tttcttgctg tacctgctgc tcttcaactg cagtggggtg gaggcagggtg 180  
agaatgcggg taaggatgca ggtaagaaaa agtgctcggg gagctcggac agcggctccg 240  
ggttctggaa ggccctgacc ttcattggccg tcggaggagg actcgcagtc gccgggctgc 300  
ccgcgctggg cttcaccggc gccggcatcg cggccaactc ggtggctgcc tcgctgatga 360  
gctggctctg gatcctgaat gggggcggcg tgcccgcggg ggggctagtg gccacgctgc 420  
agagcctcgg ggctgggtggc agcagcgtcg tcataggtaa tattgggtgcc ctgatgggct 480  
acgccacca caagtatctc gatagtgagg aggatgagga gtagccagca gctcccagaa 540

cctcttcttc	cttcttggcc	taactcttcc	agttaggatc	tagaactttg	cctttttttt	600
tttttttttt	tttttttgag	atgggttctc	actatatgt	ccaggctaga	gtgcagtggc	660
tattcacaga	tgcgaacata	gtacactgca	gcctccaact	cctagcctca	agtgatcctc	720
ctgtctcaac	ctcccaagta	ggattacaag	catgcgccga	cgatgccag	aatccagaac	780
tttgctatc	actctcccca	acaacctaga	tgtgaaaaca	gaataaactt	caccagaaa	840
a						841

<210> 124

<211> 4652

<212> DNA

<213> NM\_183047.1| Homo sapiens protein kinase C binding protein 1 (PRKCBP1), transcript variant 1, mRNA

<400> 124	
gtgagaacta	ggagcctgtc ctccatgttt tataagtatt gacattacac agtgттааса 60
atgcatccac	agagcttggc tgaagaggaa ataaaaacag aacaggaggt ggtagagggc 120
atggatatct	ctactcgctc caaagatcct ggctctgcag agagaacagc ccagaaaaga 180
aagttcccca	gccctccaca ttcttccaat ggccactcgc cgcaggacac atcaacaagc 240
cccattaaaa	agaaaaagaa acctggctta ctgaacagta acaataagga gcagtcagaa 300
ctaagacatg	gtccgtttta ctatatgaag cagccactca ccacagaccc tgttgatgtt 360
gtaccgcagg	atggacggaa tgattttctac tgctgggttt gtcaccggga aggccaagtc 420
ctttgctgtg	agctctgtcc ccgggtttat cacgctaagt gtctgagact gacatcgga 480
ccagaggggg	actggttttg tcctgaatgt gagaaaatta cagtagcaga atgcatcgag 540
accagagta	aagccatgac aatgctcacc attgaacagt taccctacct gctcaagttt 600
gccattcaga	aaatgaaaca gccagggaca gatgcattcc agaagcccgt tccattggaa 660
cagcaccctg	actatgcgga atacatcttc catccaatgg acctttgtac attggaaaag 720
aatgcgaaaa	agaaaatgta tggctgcaca gaagccttcc tggctgatgc aaagtggatt 780
ttgcacaact	gcatcattta taatggggga aatcacaaat tgacgcaaат agcgaaagta 840
gtcatcaaaa	tctgtgaaca tgagatgaat gaaatcgaag tatgtccaga atgttatcta 900
gctgcttgcc	aaaaacgaga taactggttt tgtgagcctt gtagcaatcc acatcctttg 960
gtctggggcca	aactgaaggg gtttccattc tggcctgcaa aagctctaag ggataaagac 1020
gggcaggctg	atgcccгatt ctttgгacaa catgacaggg cctgggttcc aataaataat 1080
tgctacctca	tgtctaaaga aattcctttt tctgtgaaaa agactaagag catcttcaac 1140
agtgccatgc	aagagatgga ggtttacgtg gagaacatcc gcaggaagtt tggggttttt 1200
aattactctc	catttaggac accctacaca cccaacagcc agtatcaaат gctgctcgat 1260

cccaccaacc	ccagcgccgg	cactgccaaag	atagacaagc	aggagaagggt	caagctcaac	1320
tttgacatga	cggcatcccc	caagatcctg	atgagcaagc	ctgtgctgag	tgggggcaca	1380
ggccgcccga	tttccttgct	ggatatgccg	cgctccccc	tgagcacaaa	ctcttctgtg	1440
cacacgggct	ccgacgtgga	gcaggatgct	gagaagaagg	ccacgtcgag	ccacttcagt	1500
gcgagcgagg	agtccatgga	cttcctggat	aagagcacag	cttcaccagc	ctccaccaag	1560
acggggacaag	cagggagttt	atccggcagc	ccaaagccct	tctctcctca	actgtcagct	1620
cctatcacga	cgaaaacgga	caaaacctcc	accaccggca	gcacccctgaa	tcttaacctg	1680
gatcgaagca	aagctgagat	ggatttgaag	gagctgagcg	agtcggtcca	gcaacagtcc	1740
acccctgttc	ctctcatctc	tcccaagcgc	cagattcgta	gcaggttcca	gctgaatctt	1800
gacaagacca	tagagagttg	caaagcacaa	ttaggcataa	atgaaatctc	ggaagatgtc	1860
tatacggccg	tagagcacag	cgattcggag	gattctgaga	agtcagatag	tagcgatagt	1920
gagtatatca	gtgatgatga	gcagaagtct	aagaacgagc	cagaagacac	agaggacaaa	1980
gaaggttgtc	agatggacaa	agagccatct	gctgttaaaa	aaaagcccaa	gcctacaaac	2040
ccagtggaga	ttaaagagga	gctgaaaagc	acgtcaccag	ccagcgagaa	ggcagaccct	2100
ggagcagtca	aggacaaggc	cagccctgag	cctgagaagg	acttttccga	aaaggcaaaa	2160
ccttcacctc	accccataaa	ggataaactg	aagggaaaag	atgagacgga	ttccccaaca	2220
gtccatttgg	gcctggactc	tgattcagag	agcgaacttg	tcatagattt	aggagaagac	2280
cattctgggc	gggaggggtc	aaaaaataag	aaggaaccca	aagaaccatc	tcccaaacag	2340
gatgttgtag	gtaaaactcc	accatccacg	acggtgggca	gccattctcc	cccggaaaaca	2400
ccggtgctca	cccgtcttcc	cgcctcaact	tccgcggctg	gcgccacagc	caccaccagc	2460
acgtcctcca	cggtcaccgt	cacggccccg	gcccccgccg	ccacaggaag	cccagtgaaa	2520
aagcagaggc	cgtttttacc	gaaggagact	gccccggccg	tgacgcgggt	cgtgtggaac	2580
tcacaaactg	tccagcagaa	ggagatcaca	cagagcccat	ccacgtccac	catcacccctg	2640
gtgaccagca	cacagtcac	gccccctggc	accagctcgg	ggccatgag	cacccttgtg	2700
tcctcagtca	acgtgacct	gcccacgccc	actgcctcag	ctgatgtcgc	cgctgatatt	2760
gccaagtaca	ctagcaaaat	gatggatgca	ataaaaggaa	caatgacaga	aatatacaac	2820
gatctttcta	aaaacactac	tggaagcaca	atagctgaga	ttcgagggt	gaggatcgag	2880
atagagaagc	tccagtggct	gcaccagcaa	gagctctccg	aaatgaaaca	caacttagag	2940
ctgaccatgg	cggagatgag	gcagagcctg	gagcaggagc	gggaccgggt	catcgccgag	3000
gtgaagaagc	agctggagtt	ggagaagcag	caggcgggtg	atgagaccaa	gaagaagcag	3060
tgggtgcgca	actgcaagaa	ggaggccatc	ttttactgct	gttggaaac	tagctactgt	3120
gactaccct	gccagcaagc	ccactggcct	gagcacatga	agtcctgcac	ccagtcagct	3180
actgctcctc	agcaggaagc	ggatgctgag	gtgaacacag	aaacactaaa	taagtcctcc	3240
caggggagct	cctcgagcac	acaatcagca	ccttcagaaa	cggccagcgc	ctccaaagag	3300

```

aaggagacgt cagctgagaa aagcaaggag agtggctcga cccttgacct ttctggctcc 3360
agagagacgc cctcctccat tctcttaggc tccaaccaag gctctgttag caaaagggtgt 3420
gacaagcaac ctgcctatgc cccaaccacc acagaccacc agccgcaccc caactacccc 3480
gcccagaagt accattcccc gagtaataaa tccagttgga gcagcagtga tgagaagagg 3540
ggatcgacac gttccgatca caacaccagt accagcacga agagcctcct ccgaaaagag 3600
tctcggctgg acaccttctg ggactagcag tgaatcgga cacaaccac ccacccatt 3660
gggagaaaaa cccagacgcc aggaaaagaa gaaacaaca aggcaggaga acagccactt 3720
tcagacttga aaatgacaaa accctcagtt gagcctgagc ccccggcgcg ggggctgcta 3780
cactacagga caccagcat cggctttgac tgcagactgt tcaccacac gagccctgtg 3840
cttttggtgt aaataatgta caatttggtg atgtcattga atctagagga ctttccctt 3900
tttatatttg tattaacttt aacttattaa aaaaaaaaaa agaaaaagaa aaacgattta 3960
aaaaaaaaa aaaaagcaac caacccaac aacaaaaaag aatgttttgg tattggagaa 4020
gggatggtca gttagcctgt ctgtcacacg acggaatgga tactgggccc ggggaccact 4080
ttcatactca cgtcctcatc cttggatacc caggggaggg cgaaccgttt tcgctcgtgt 4140
gtctgtacgc agcatgttgg gatcgggagt ttcggcacag actatcccat caagccgttg 4200
gctcctttca gctactacgt taccacgttc ctaaaacgca agctctccgg accagacgga 4260
cacagggaga agctagtttc tttcatgtga ttgaaatgat gactctactc ctaaaaggga 4320
aaaaacaata tccttgttta cagaagagaa acaacaagc cccactcagc tcagtcacag 4380
gagagaacac agaaagtctt aggatcatga actctgaaaa aaagagaaac cttatctttg 4440
ctttgtggtt cttttaacaa cactcacaca cacttggtca gagatgctgt gcttcttgga 4500
agcaaggact caaaggcaag gtgcacgcag aggacgtttg agtctgggat gaagcatgta 4560
cgtattatth atatgatgga atttcacgtt tttatgtaag catgaaacac aggcagtatg 4620
agagaaagca aggcccgta tgctgtccgt ac 4652

```

<210> 125

<211> 3217

<212> DNA

<213> NM\_017452.1| Homo sapiens staufer, RNA binding protein (Drosophila) (STAU), transcript variant T2, mRNA

<400> 125

```

acttcctgcc gggctgcggg cgcctgagcg ctcttcagcg ttgcgcggc ggctgcgcgt 60
ctctctcggc tcccgcttcc ttgaccgcc tcccccccc ggcccggcgg cgccgcctc 120
ctccacggcc actccgcctc ttccctccct tcgtcccttc ttctctccc tttttcctt 180
cttccttccc ctctcgccg ccaccgccca ggaccgccgg ccgggggacg agtccggagc 240

```

agcagccaga gtttattaac cacttaacct ctcagaactg aacaaagaca acattgttcc	300
tggaacgccc tctttttaaa aaagaaagca taacccttac tgtagaacta aatgcactgt	360
gcatgaaact tggaaaaaaa ccaatgtata agcctgttga cccttactct cggatgcagt	420
ccacctataa ctacaacatg agaggagggtg cttatccccc gaggtacttt taccatttc	480
cagttccacc ttacttttat caagtggaac tttctgtggg aggacagcaa tttaatggca	540
aaggaaagac aagacaggct gcgaaacacg atgctgctgc caaagcggtg aggatcctgc	600
agaatgagcc cctgccagag aggctggagg tgaatggaag agaatccgaa gaagaaaatc	660
tcaataaatc tgaaataagt caagtgtttg agattgcact taaacggaac ttgcctgtga	720
atttcgaggt ggcccgggag agtggcccac ccacatgaa gaactttgtg accaaggttt	780
cggttgggga gtttgtgggg gaaggtgaag ggaaaagcaa gaagatttca aagaaaaatg	840
ccgccatagc tgttcttgag gagctgaaga agttaccgcc cctgcctgca gttgaacgag	900
taaagcctag aatcaaaaag aaaacaaaac ccatagtcaa gccacagaca agcccagaat	960
atggccaggg gatcaatccg attagccgac tggccagat ccagcaggca aaaaaggaga	1020
aggagccaga gtacacgctc ctcacagagc gaggcctccc gcgccgcagg gagtttgtga	1080
tgcaggtgaa ggttggaac cactctgag aaggaaacggg caccaacaag aaggtggcca	1140
agcgcaatgc agccgagaac atgctggaga tccttggttt caaagtcccg cagcggcagc	1200
ccaccaaac ccgactcaag tcagaggaga agacacccat aaagaaacca ggggatggaa	1260
gaaaagtaac cttttttgaa cctggctctg gggatgaaaa tgggactagt aataaagagg	1320
atgagttcag gatgccttat ctaagtcac agcagctgcc tgctggaatt cttcccatgg	1380
tgcccagagt cgcccaggct gtaggagtta gtcaaggaca tcacaccaa gattttacca	1440
gggcagctcc gaatcctgcc aaggccacgg taactgccat gatagcccga gagttgttgt	1500
atgggggcac ctgcccaca gccgagacca ttttaaagaa taacatctct tcaggccacg	1560
taccccatgg acctctcac agaccctctg agcaactgga ctatctttcc agagtccagg	1620
gattccaggt tgaatacaaa gacttcccca aaaacaacaa gaacgaattt gtatctctta	1680
tcaattgctc ctctcagcca cctctgatca gccatggtat cggcaaggat gtggagtcct	1740
gccatgatat ggctgcgctg aacatcttaa agttgctgtc tgagttggac caacaaagta	1800
cagagatgcc aagaacagga aacggacca tgtctgtgtg tgggaggtgc tgaacctttt	1860
ctggccatga accattataa aatcccaaca tatatactga aaatactgaa actgctttga	1920
aaatttgga tttctgata ctccagtggg ccgagagaca cggtaggtta aggatgtggg	1980
cagcagcagg gaagacaaca gaaacacaag gaggcggctg tggccggctg gactgtgctg	2040
gggtttgttg tgatggccac tcggtgacct ggcggtccct acgcaatagc agctgcctgt	2100
ggggaagaag ggctgcccag ccagctgggt ctcgccgggac accagcagat ccacacctg	2160
ggcacctccg tgtttggctt ttttttccc ctgtgtgaaa gaagaaacgg cacgacctt	2220

tctcaagctg gctcactcag acacattggg acaaaccctg gacagccatg ccagagagag	2280
gcctttgacc ggccccagag ctaaaagcac cagagaaaat caaatgcttc ctactcagcg	2340
tgacccaact tttctagtgt gccacggccc caccacctcc tgcagtaccc acaccatcac	2400
cactgctttc tcttccaaca gtgatctgta ttcttagttt cattattttc ttttgattga	2460
tatgacacta tataaaatTT tcatttgaga atttctcaat tgtatctagt taaatagcac	2520
agtttggaag cttgtctgag actgacttta tcaataatct aaccgacaaa gatcatatcc	2580
atgtgtatgt ggtttagacat ttttatttca ttgactaacc caggacagtt tcagtgatgc	2640
aaattgtgtg ccctctgggt cagctgaaac agtcctggac tttcaaaaac cttgaataag	2700
tctcccacag ttgtataaat tggacaattt aggaatttta aacttttagat gatcatttgg	2760
ttccattttt atttcatttt tatttttggt aatgcaaaca ggacttaaat gaactttgat	2820
ctctgtttta aagattatta aaaaacattg tgtatctata catatggctc ttgaggactt	2880
agctttcact acactacagg atatgatctc catgtagtcc atataaacct gcagagtgat	2940
ttccagaggt gctcgatact gttaattaca tctccattag ggctgaaaag aatgacctac	3000
gtttctgtat acagctgtgt tgcttttgat gttgtgttac tgtacacaga agtgtgtgca	3060
ctgaggctct gcgtgtggtc cgtatggaaa acctggtagc cctgcgagtt aagtactgct	3120
tccattcatt gtttacgctg gaatttttct ccccatggaa tgtaagtaaa acttaagtgt	3180
ttgtcatcaa taaatggtaa tactaaaaaa aaaaaaa	3217

<210> 126

<211> 3506

<212> DNA

<213> NM\_017453.1| Homo sapiens staufen, RNA binding protein (Drosophila) (STAU), transcript variant T3, mRNA

<400> 126	
acttcctgcc gggctgcggg cgcttgagcg ctcttcagcg tttgcgcggc ggctgcgcgt	60
ctctctcggc tcccgtttcc tttgaccgcc tccccccccc ggcccggcgg cgcccgcctc	120
ctccacggcc actccgcctc ttccctccct tcgtcccttc ttctctccc ttttttcctt	180
cttccttccc ctctcgcgg ccaccgccc ggaccgccc cggggggacg agtccggagc	240
agcagccaga gtttattaac cacttaacct ctcagaactg aacaaagaca acattgttcc	300
tggaacgccc tctttttaaa aaaggtagaa cttagactt catagcactg aattaacctg	360
cactgaaagc tgttttacct catttggtca cttttgttga aagtgaccat gtctcaagtt	420
caagtgaag ttcagaacct atctgtgct ctctcagga gccaaatact gaacaagaac	480
cagtctcttc tctcacagcc tttgatgagt attccttcta ctactagctc tctgccctct	540
gaaaatgcag gtagacccat tcaaaactct gctttaccct ctgcatctat tacatccacc	600

agtgacgctg	cagaaagcat	aacccctact	gtagaactaa	atgcactgtg	catgaaactt	660
ggaaaaaaac	caatgtataa	gcctgttgac	ccttactctc	ggatgcagtc	cacctataac	720
tacaacatga	gaggaggtgc	ttatcccccg	aggtactttt	acccatttcc	agttccacct	780
ttactttatc	aagtggaaact	ttctgtggga	ggacagcaat	ttaatggcaa	aggaaagaca	840
agacaggctg	cgaaacacga	tgctgctgcc	aaagcgttga	ggatcctgca	gaatgagccc	900
ctgccagaga	ggctggaggt	gaatgggaaga	gaatccgaag	aagaaaatct	caataaatct	960
gaaataagtc	aagtgtttga	gattgcactt	aaacggaact	tgctgtgaa	tttcgaggtg	1020
gccccgggaga	gtggccccacc	ccacatgaag	aactttgtga	ccaaggtttc	ggttggggag	1080
tttgtggggg	aaggtgaagg	gaaaagcaag	aagatttcaa	agaaaaatgc	cgccatagct	1140
gttcttgagg	agctgaagaa	gttaccgccc	ctgcctgcag	ttgaacgagt	aaagcctaga	1200
atcaaaaaga	aaacaaaacc	catagtcaag	ccacagacaa	gcccagaata	tggccagggg	1260
atcaatccga	ttagccgact	ggcccagatc	cagcaggcaa	aaaaggagaa	ggagccagag	1320
tacacgctcc	tcacagagcg	aggcctcccc	cgccgcaggg	agtttgtgat	gcaggtgaag	1380
gttggaacc	acactgcaga	aggaacgggc	accaacaaga	aggtggccaa	gcgcaatgca	1440
gccgagaaca	tgctggagat	ccttggtttc	aaagtccgc	agcggcagcc	caccaaacc	1500
gcactcaagt	cagaggagaa	gacaccata	aagaaaccag	gggatggaag	aaaagtaacc	1560
ttttttgaac	ctggctctgg	ggatgaaaat	gggactagta	ataaagagga	tgagttcagg	1620
atgccttate	taagtcatca	gcagctgcct	gctggaattc	ttcccatggg	gcccaggtc	1680
gcccaggctg	taggagttag	tcaaggacat	cacaccaaag	attttaccag	ggcagctccg	1740
aatcctgcc	aggccacggg	aactgccatg	atagcccag	agttgttgta	tgggggcacc	1800
tcgcccacag	ccgagaccat	tttaaagaat	aacatctctt	caggccacgt	accccatgga	1860
cctctcacga	gaccctctga	gcaactggac	tatctttcca	gagtccaggg	attccaggtt	1920
gaatacaaag	acttcccca	aaacaacaag	aacgaatttg	tatctcttat	caattgctcc	1980
tctcagccac	ctctgatcag	ccatggtatc	ggcaaggatg	tggagtcttg	ccatgatatg	2040
gctgcgctga	acatcttaaa	gttgctgtct	gagttggacc	aacaaagtac	agagatgcca	2100
agaacaggaa	acggaccaat	gtctgtgtgt	gggaggtgct	gaaccttttc	tggccatgaa	2160
ccattataaa	atcccaacat	atatactgaa	aatactgaaa	ctgctttgaa	aatttggaat	2220
ttctgatacc	tccagtgggc	cgagagacac	ggtgggtaaa	ggatgtgggc	agcagcaggg	2280
aagacaacag	aaacacaagg	aggcggctgt	ggccggctgg	actgtgctgg	ggtttgttgt	2340
gatggccact	cggtgacctg	gcggtcccta	cgcaatagca	gctgcctgtg	gggaagaagg	2400
gctgcccagc	cagctggttc	ttccgggaca	ccagcagatc	cacaccctgg	gcacctccgt	2460
gtttggtctt	ttttttcccc	tgtgtgaaag	aagaaacggc	acgaccctt	ctcaagctgg	2520
ctcactcaga	cacattggga	caaaccctgg	acagccatgc	cagagagagg	cctttgaccg	2580
gccccagagc	taaaagcacc	agagaaaatc	aatgtcttcc	tactcagcgt	gacccaactt	2640

ttctagtgtg ccacggcccc accacctcct gcagtaccca caccatcacc actgctttct	2700
cttccaacag tgatctgtat tcttagtttc attattttct tttgattgat atgacactat	2760
ataaaatttt catttgagaa tttctcaatt gtatctagtt aaatagcaca gtttggaac	2820
ttgtctgaga ctgactttat caataatcta accgacaaag atcatatcca tgtgtatgtg	2880
gtagacatt tttatttcat tgactaacc aggacagttt cagtgatgca aattgtgtgc	2940
cctctgggtc agctgaaaca gtcctggact ttcaaaaacc ttgaataagt ctcccacagt	3000
tgtataaatt ggacaattta ggaattttta acttttagatg atcatttggg tccattttta	3060
tttcattttt atttttgtta atgcaaacag gacttaaagt aactttgatc tctgttttaa	3120
agattattaa aaaacattgt gtatctatac atatggctct tgaggactta gctttcacta	3180
cactacagga tatgatctcc atgtagtcca tataaacctg cagagtgatt ttccagagtg	3240
ctcgatactg ttaattacat ctccattagg gctgaaaaga atgacctacg tttctgtata	3300
cagctgtgtt gcttttgatg ttgtgttact gtacacagaa gtgtgtgcac tgaggctctg	3360
cgtgtggtcc gtatggaaaa cctggtagcc ctgagagtta agtactgctt ccattcattg	3420
tttacgctgg aatttttctc cccatggaat gtaagtaaaa cttagtgtt tgtcatcaat	3480
aatggtaat actaaaaaaaa aaaaaa	3506

<210> 127

<211> 4538

<212> DNA

<213> NM\_199169.1| Homo sapiens transmembrane, prostate androgen induced RNA (TMEPAI), transcript variant 2, mRNA

<400> 127	
tggtcgtcct ccttgggttc ggggtgaaagc gcttgggggt tcagtgggccc atgatccccg	60
agctgctgga gaactgaagg cggacagtct cctgcgaaac caggcaatgg cggagctgga	120
gtttgttcag atcatcatca tcgtggtggt gatgatggtg atggtggtgg tgatcacgtg	180
cctgctgagc cactacaagc tgtctgcacg gtccttcac agccggcaca gccaggggagc	240
gaggagagaa gatgccctgt cctcagaagg atgcctgtgg ccctcggaga gcacagtgtc	300
aggcaacgga atcccagagc cgcaggtcta cgccccgcct cggcccaccg accgcctggc	360
cgtgccgccc ttgcgccagc gggagcgctt ccaccgcttc cagcccacct atccgtacct	420
gcagcacgag atcgacctgc caccaccat ctcgctgtca gacggggagg agccccacc	480
ctaccagggc ccctgcaccc tccagcttcg ggaccccgag cagcagctgg aactgaaccg	540
ggagtcggtg cgcgcacccc caaacagaac catcttcgac agtgacctga tggatagtgc	600
caggctgggc ggccccctgcc ccccagcag taactcgggc atcagcgcca cgtgctacgg	660
cagcggcggg cgcatggagg ggccgcccgc cacctacagc gaggtcatcg gccactaccc	720



ggggtcctcc	ttccagcacc	agcagagcag	tggggccgccc	tccttgctgg	aggggacccg	780
gctccaccac	acacacatcg	cgcccctaga	gagcgcagcc	atctggagca	aagagaagga	840
taaacagaaa	ggacaccctc	tctaggggtcc	ccagggggggc	cgggctgggg	ctgcgtaggt	900
gaaaaggcag	aacactccgc	gcttcttaga	agaggagtga	gaggaaggcg	gggggcgag	960
caacgcacg	tgtggccctc	ccctcccacc	tccctgtgta	taaatattta	catgtgatgt	1020
ctggtctgaa	tgcacaagct	aagagagctt	gcaaaaaaaaa	aaagaaaaaa	gaaaaaaaaa	1080
aaccacgttt	ctttgttgag	ctgtgtcttg	aaggcaaaag	aaaaaaaatt	tctacagtag	1140
tctttcttgt	ttctagttga	gctgcgtgcg	tgaatgctta	ttttcttttg	tttatgataa	1200
tttcacttaa	ctttaagac	atatttgcac	aaaacctttg	tttaaagatc	tgcaatatta	1260
tatatataaa	tatatataag	ataagagaaa	ctgtatgtgc	gagggcagga	gtatttttgt	1320
attagaagag	gcctattaaa	aaaaaaagtt	gttttctgaa	ctagaagagg	aaaaaaatgg	1380
caatttttga	gtgccaaagtc	agaaagtgtg	tattaccttg	taaagaaaaa	aattacaaag	1440
caggggttta	gagttattta	tataaatgtt	gagattttgc	actatTTTTT	aatataaata	1500
tgtcagtgtc	tgcttgatgg	aaacttctct	tgtgtctgtt	gagactttta	gggagaaatg	1560
tcggaatttc	agagtcgcct	gacggcagag	ggtgagcccc	cgtggagtct	gcagagaggc	1620
cttggccagg	agcggcgggc	tttcccagg	ggccactgtc	cctgcagagt	ggatgcttct	1680
gcctagtgtg	aggttatcac	cacgttatat	attccctacc	gaaggagaca	ccttttcccc	1740
cctgaccag	aacagccttt	aaatcacaa	caaaatagga	aagttaacca	cggaggcacc	1800
gagttccagg	tagtggtttt	gcctttccca	aaaatgaaaa	taaactgtta	ccgaagggaat	1860
tagtttttcc	tcttcttttt	tccaactgtg	aagggtcccc	tggggtggag	catggtgccc	1920
ctcacaagcc	gcagcggctg	gtgcccgggc	taccaggggac	atgccagagg	gctcgatgac	1980
ttgtctctgc	agggcgcttt	ggtggttggt	cagctggcta	aagggttcacc	ggtgaaggca	2040
ggtgcggtaa	ctgccgcact	ggaccctagg	aagccccagg	tattcgcaat	ctgacctcct	2100
cctgtctgtt	tcccttcacg	gatcaattct	cacttaagag	gccaataaac	aacccaacat	2160
gaaaagggtga	caagcctggg	tttctcccag	gatagggtgaa	aggggttaaaa	tgagtaaagc	2220
agttgagcaa	acaccaaccc	gagcttcggg	cgcagaattc	ttcaccttct	cttccccctt	2280
ccatctcctt	ttcccgcgga	aacaacgctt	cccttctggt	gtgtctgttg	atctgtgttt	2340
tcattttacat	ctctcttaga	ctccgctctt	gttctccagg	ttttcaccag	atagattttg	2400
ggttggcggg	acctgctggt	gacgtgcagg	tgaaggacag	gaaggggcat	gtgagcgtaa	2460
atagagggtga	ccagaggaga	gcatgagggg	tggggctttg	ggaccacccg	gggccagtg	2520
ctggagcttg	acgtctttcc	tccccatggg	ggtggggagg	ccccagctg	gaagagcaga	2580
ctcccagctg	ctacccccct	ccttcccctg	ggagtggctt	tccatttttg	gcagaatgct	2640
gactagtaga	ctaacataaa	agatataaaa	ggcaataact	attgtttgtg	agcaactttt	2700

ttataacttc	caaaacaaaa	acctgagcac	agttttgaag	ttctagccac	tcgagctcat	2760
gcatgtgaaa	cgtgtgcttt	acgaagggtg	cagctgacag	acgtgggctc	tgcattgccgc	2820
cagcctagta	gaaagtcttc	gttcattggc	aacagcagaa	cctgcctctc	cgtgaagtcg	2880
tcagcctaaa	atttgtttct	ctcttgaaga	ggattctttg	aaaaggctct	gcagagaaat	2940
cagtacaggt	tatccccgaa	ggtacaagga	cgcacttgta	aagatgatta	aaacgtatct	3000
ttcctttatg	tgacgcgtct	ctagtgcctt	actgaagaag	cagtgcacct	cccgtcgctc	3060
ggtgaggacg	ttccccggac	gtgcctcact	cacctgggac	tggatatccc	tcccagggctc	3120
caccaagggc	tcctgctttt	cagacacccc	atcatcctcg	cgcgtcctca	ccctgtctct	3180
accagggagg	tgcctagctt	ggtgagggtta	ctcctgctcc	tccaaccttt	ttttgccaa	3240
gtttgtacac	gactcccatc	taggctgaaa	acctagaagt	ggaccttggt	tgtgtgcatg	3300
gtgtcagccc	aaagccaggc	tgagacagtc	ctcatatcct	cttgagccaa	actgtttggg	3360
tctcgttgct	tcatgggatg	gtctggattt	gtgggaatgg	ctttgcgtga	gaaaggggag	3420
gagagtgggt	gctgccttca	gccggcttga	ggacagagcc	tgtccctctc	atgacaactc	3480
agtgttgaag	cccagtgtcc	tcagcttcat	gtccagtggg	tggcagaagt	tcatggggta	3540
gtggcctctc	aaaggctggg	cgcattccca	gacagccagc	aggttgtctc	tggaaacgac	3600
cagagttaag	ctctcggctt	ctctgctgag	ggtgcacctt	ttcctctaga	tggtagttgt	3660
cacgttatct	ttgaaaactc	ttggactgct	cctgaggagg	ccctcttttc	cagtaggaag	3720
ttagatgggg	gttctcagaa	gtggctgatt	ggaaggggac	aagcttcggt	tcaggggtct	3780
gccgttccat	cctgggttcag	agaaggccga	gcgtggcttt	ctctagcctt	gtcactgtct	3840
ccctgcctgt	caatcaccac	ctttcctcca	gaggaggaaa	attatctccc	ctgcaaagcc	3900
cggttctaca	cagatttcac	aaattgtgct	aagaaccgtc	cgtgttctca	gaaagcccag	3960
tgtttttgca	aagaatgaaa	agggacccca	tatgtagcaa	aaatcagggc	tgggggagag	4020
ccgggttcat	tccctgtcct	cattgggtcgt	ccctatgaat	tgtacgtttc	agagaaattt	4080
tttttcctat	gtgcaacacg	aagcttccag	aaccataaaa	tatcccgtcg	ataaggaaag	4140
aaaatgtcgt	tgttggttgt	tttctggaaa	ctgcttgaaa	tcttgctgta	ctatagagct	4200
cagaaggaca	cagcccgtcc	tcccctgcct	gcctgattcc	atggctgttg	tgctgattcc	4260
aatgctttca	cgttgggttc	tggcgtggga	actgctctcc	tttgcagccc	catttcccaa	4320
gctctgttca	agttaaactt	atgtaagctt	tccgtggcat	gcggggcgcg	caccacgctc	4380
cccgtgcgt	aagactctgt	atttggatgc	caatccacag	gcctgaagaa	actgcttggt	4440
gtgtatcagt	aatcattagt	ggcaatgatg	acattctgaa	aagctgcaat	acttatacaa	4500
taaattttac	aattcttttg	aaaaaaaaaa	aaaaaaaaaa			4538

<210> 128

<211> 4531

<212> DNA

<213> NM\_199170.1| Homo sapiens transmembrane, prostate androgen induced RNA (TMEPAI), transcript variant 3, mRNA

```
<400> 128
tggtcgtcct ccttgggttc gggtgaaagc gcttgggggt tcagtgggcc atgatccccg    60
agctgctgga gaactgaagg cggacagtct cctgcgaaac cagcggagct ggagtttgtt    120
cagatcatca tcatcgtggt ggtgatgatg gtgatggtgg tggtgatcac gtgcctgctg    180
agccactaca agctgtctgc acggtccttc atcagccggc acagccaggg gcggaggaga    240
gaagatgccc tgtcctcaga aggatgcctg tggccctcgg agagcacagt gtcaggcaac    300
ggaatcccag agccgcaggt ctacgccccg cctcggccca ccgaccgcct ggccgtgccg    360
cccttcgccc agcgggagcg cttccaccgc ttccagccca cctatccgta cctgcagcac    420
gagatcgacc tgccaccac catctcgctg tcagacgggg aggagcccc accctaccag    480
ggcccctgca ccctccagct tcgggacccc gagcagcagc tggaaactgaa ccgggagtcg    540
gtgcgcgcac ccccaaacag aaccatcttc gacagtgacc tgatggatag tgccaggctg    600
ggcggccccct gccccccag cagtaactcg ggcacgagc ccacgtgcta cggcagcggc    660
gggcgcatgg agggggccgc gccacctac agcgagggtca tcggccacta cccgggggtcc    720
tccttcagc accagcagag cagtggggcg ccttccttgc tggagggggac ccggctccac    780
cacacacaca tcgcgcccct agagagcgca gccatctgga gcaaagagaa ggataaacag    840
aaaggacacc ctctctaggg tccccagggg ggccgggctg gggctgcgta ggtgaaaagg    900
cagaacactc cgcgcttctt agaagaggag tgagaggaag gcggggggcg cagcaacgca    960
tcgtgtggcc ctcccctccc acctccctgt gtataaatat ttacatgtga tgtctggtct   1020
gaatgcacaa gctaagagag cttgcaaaaa aaaaaagaaa aaagaaaaaa aaaaaccacg   1080
tttctttgtt gagctgtgtc ttgaaggcaa aagaaaaaaa atttctacag tagtctttct   1140
tgtttctagt tgagctgcgt gcgtgaatgc ttattttctt ttgtttatga taatttcact   1200
taactttaaa gacatatttg cacaaaacct ttgtttaaag atctgcaata ttatatatat   1260
aaatatatat aagataagag aaactgtatg tgcgagggca ggagtatttt tgtattagaa   1320
gaggcctatt aaaaaaaaaa gttgttttct gaactagaag aggaaaaaaaaa tggcaatttt   1380
tgagtgccaa gtcagaaagt gtgtattacc ttgtaaagaa aaaaattaca aagcaggggt   1440
ttagagttat ttatataaat gttgagattt tgcactattt tttaataata atatgtcagt   1500
gcttgcttga tggaaacttc tcttgtgtct gttgagactt taaggagaga atgtcggaat   1560
ttcagagtcg cctgacggca gagggtgagc ccccgaggag tctgcagaga ggccttggcc   1620
aggagcggcg ggctttcccg aggggccact gtccctgcag agtggatgct tctgcctagt   1680
gacaggttat caccacgtta tatattccct accgaaggag acaccttttc cccctgacc   1740
cagaacagcc tttaaatac aagcaaaata ggaaagttaa ccacggaggc accgagttcc   1800
```

aggtagtgggt	tttgcctttc	ccaaaaatga	aaataaactg	ttaccgaagg	aattagtttt	1860
tcctcttctt	ttttccaact	gtgaagggtcc	ccgtgggggtg	gagcatgggtg	cccctcacia	1920
gccgcagcgg	ctgggtgccc	ggctaccagg	gacatgccag	agggctcgat	gacttgtctc	1980
tgcagggcgc	tttgggtggtt	gttcagctgg	ctaaagggttc	accggtgaag	gcaggtgcgg	2040
taactgccgc	actggaccct	aggaagcccc	aggtattcgc	aatctgacct	cctcctgtct	2100
gtttcccttc	acggatcaat	tctcacttaa	gaggccaata	aacaacccaa	catgaaaagg	2160
tgacaagcct	gggtttctcc	caggataggt	gaaaggggtta	aaatgagtaa	agcagttgag	2220
caaacaccaa	cccgaacttc	gggcgcagaa	ttcttcacct	tctcttcccc	ttccatctc	2280
ctttccccgc	ggaaacaacg	cttcccttct	ggtgtgtctg	ttgatctgtg	ttttcattta	2340
catctctctt	agactccgct	cttgttctcc	aggttttcac	cagatagatt	tggggttggc	2400
gggacctgct	ggtgacgtgc	aggtgaagga	caggaagggg	catgtgagcg	taaatagagg	2460
tgaccagagg	agagcatgag	gggtggggct	ttgggaccca	ccggggccag	tggctggagc	2520
ttgacgtctt	tcctccccat	gggggtggga	gggccccag	ctggaagagc	agactcccag	2580
ctgctacccc	ctcccttccc	atgggagtgg	ctttccattt	tgggcagaat	gctgactagt	2640
agactaacat	aaaagatata	aaaggcaata	actattgttt	gtgagcaact	tttttataac	2700
ttccaaaaca	aaaacctgag	cacagttttg	aagttctagc	cactcgagct	catgcatgtg	2760
aaacgtgtgc	tttacgaagg	tggcagctga	cagacgtggg	ctctgcatgc	cgccagccta	2820
gtagaaagtt	ctcgttcatt	ggcaacagca	gaacctgcct	ctccgtgaag	tcgtcagcct	2880
aaaatttggt	tctctcttga	agaggattct	ttgaaaagg	cctgcagaga	aatcagtaca	2940
ggttatcccg	aaagggtacaa	ggacgcactt	gtaaagatga	ttaaaacgta	tctttccttt	3000
atgtgacgcg	tctctagtgc	cttactgaag	aagcagtgac	actcccgtcg	ctcggtgagg	3060
acgttcccg	acagtgcctc	actcacctgg	gactgggtatc	ccctcccagg	gtccaccaag	3120
ggctcctgct	tttcagacac	cccatcatcc	tcgcgcgtcc	tcaccctgtc	tctaccagg	3180
aggtgcctag	cttgggtgagg	ttactcctgc	tcctccaacc	tttttttgcc	aagggttgta	3240
cacgactccc	atctaggctg	aaaacctaga	agtggacctt	gtgtgtgtgc	atggtgtcag	3300
cccaaagcca	ggctgagaca	gtcctcatat	cctcttgagc	caaactgttt	gggtctcggt	3360
gcttcatgg	atgggtctgga	tttgtgggaa	tggctttgcg	tgagaaaagg	gaggagagt	3420
gttgctgccc	tcagccggct	tgaggacaga	gcctgtccct	ctcatgacaa	ctcagtgttg	3480
aagcccagtg	tcctcagctt	catgtccagt	ggatggcaga	agttcatggg	gtagtggcct	3540
ctcaaaggct	gggcgcaccc	caagacagcc	agcaggttgt	ctctggaaac	gaccagagtt	3600
aagctctcgg	cttctctgct	gaggggtgcac	cctttcctct	agatggtagt	tgtcacgtta	3660
tctttgaaaa	ctcttggaact	gctcctgagg	aggccctctt	ttccagtagg	aagttagatg	3720
ggggttctca	gaagtggctg	attggaaagg	gacaagcttc	gtttcagggg	tctgccgttc	3780

catcctggtt	cagagaaggc	cgagcgtggc	tttctctagc	cttgctactg	tctccctgcc	3840
tgtcaatcac	cacctttcct	ccagaggagg	aaaattatct	cccctgcaaa	gcccggttct	3900
acacagattt	cacaaattgt	gctaagaacc	gtccgtgttc	tcagaaagcc	cagtgttttt	3960
gcaaagaatg	aaaagggacc	ccatatgtag	caaaaatcag	ggctggggga	gagccgggtt	4020
cattccctgt	cctcattggg	cgtccctatg	aattgtacgt	ttcagagaaa	ttttttttcc	4080
tatgtgcaac	acgaagcttc	cagaaccata	aaatatcccc	tcgataagga	aagaaaatgt	4140
cgttggtgtt	gtttttctgg	aaactgcttg	aaatcttgct	gtactataga	gctcagaagg	4200
acacagcccc	tcctcccctg	cctgcctgat	tccatggctg	ttgtgctgat	tccaatgctt	4260
tcacgttggg	tcctggcgtg	ggaactgctc	tcctttgcag	ccccattttc	caagctctgt	4320
tcaagttaaa	cttatgtaag	ctttccgtgg	catgcggggc	gcgcacccac	gtccccgctg	4380
cgtaagactc	tgtatttgga	tgccaatcca	caggcctgaa	gaaactgctt	gttgtgtatc	4440
agtaatcatt	agtggcaatg	atgacattct	gaaaagctgc	aatacttata	caataaattt	4500
tacaattctt	tggaaaaaaa	aaaaaaaaaa	a			4531

<210> 129

<211> 2692

<212> DNA

<213> NM\_152871.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 2, mRNA

<400> 129	
cctacccgcg	cgcaggccaa gttgctgaat caatggagcc ctcccccaacc cgggcgttcc 60
ccagcgaggc	ttcctttccca tcctcctgac caccggggct tttcgtgagc tcgtctctga 120
tctcgcgcaa	gagtgcacaca cagggtgttca aagacgcttc tgggggagtga ggggaagcggg 180
ttacgagtga	cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggc 240
cagccctggc	tgcccaggcg gagctgcctc ttctcccgcg gggtgggtgga cccgctcagt 300
acggagttag	ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg 360
gaccctccta	cctctggttc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc 420
ccaagtgact	gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga 480
gactcagaac	ttggaaggcc tgcatcatga tggccaattc tgccataagc cctgtcctcc 540
aggtgaaagg	aaagctaggg actgcacagt caatggggat gaaccagact gcgtgccctg 600
ccaagaaggg	aaggagtaca cagacaaagc ccatttttct tccaaatgca gaagatgtag 660
atttgtgtgat	gaaggacatg gcttagaagt ggaaataaac tgcacccgga cccagaatac 720
caagtgcaga	tgtaaaccaa actttttttg taactctact gtatgtgaac actgtgaccc 780
ttgcaccaaa	tgtgaacatg gaatcatcaa ggaatgcaca ctcaccagca acaccaagtg 840

caaagaggaa	gtgaagagaa	aggaagtaca	gaaaacatgc	agaaagcaca	gaaaggaaaa	900
ccaaggttct	catgaatctc	caaccttaaa	tcctgaaaca	gtggcaataa	atztatctga	960
tgttgacttg	agtaaata	tcaccactat	tgctggagtc	atgacactaa	gtcaagttaa	1020
aggctttgtt	cgaagaatg	gtgtcaatga	agccaaaata	gatgagatca	agaatgacaa	1080
tgtccaagac	acagcagaac	agaaagttca	actgcttcgt	aattggcatc	aacttcatgg	1140
aaagaaagaa	gcgtatgaca	cattgattaa	agatctcaaa	aaagccaatc	tttgtactct	1200
tgcagagaaa	attcagacta	tcctcctcaa	ggacattact	agtgactcag	aaaattcaaa	1260
cttcagaaat	gaaatccaaa	gcttgggtcta	gagtgaaaaa	caacaaattc	agttctgagt	1320
atatgcaatt	agtgtttgaa	aagattctta	atagctggct	gtaaatactg	cttggttttt	1380
tactgggtac	atztatcat	ttattagcgc	tgaagagcca	acatatttgt	agatttttaa	1440
tatctcatga	ttctgcctcc	aaggatgttt	aaaatctagt	tgggaaaaca	aacttcatca	1500
agagtaaattg	cagtggcatg	ctaagtaccc	aaataggagt	gtatgcagag	gatgaaagat	1560
taagattatg	ctctggcatc	taacatatga	ttctgtagta	tgaatgtaat	cagtgtatgt	1620
tagtacaat	gtctatccac	aggctaacc	cactctatga	atcaatagaa	gaagctatga	1680
ccttttgctg	aaatatcagt	tactgaacag	gcaggccact	ttgcctctaa	attacctctg	1740
ataattctag	agattttacc	atattttctaa	actttgttta	taactctgag	aagatcatat	1800
ttatgtaaag	tatatgtatt	tgagtgcaga	atttaaataa	ggctctacct	caaagacctt	1860
tgcacagttt	attgggtgtca	tattatacaa	tatttcaatt	gtgaattcac	atagaaaaca	1920
ttaaattata	atgtttgact	attatatatg	tgtatgcatt	ttactggctc	aaaactacct	1980
acttctttct	caggcatcaa	aagcattttg	agcaggagag	tattactaga	gctttgccac	2040
ctctccattt	ttgccttggt	gctcatctta	atggcctaata	gcacccccaa	acatggaaat	2100
atcaccaaaa	aatacttaata	agtccaccaa	aaggcaagac	tgcccttaga	aattctagcc	2160
tggtttggag	atactaactg	ctctcagaga	aagtagcttt	gtgacatgtc	atgaacccat	2220
gtttgcaatc	aaagatgata	aaatagattc	ttatttttcc	cccacccccg	aaaatgttca	2280
ataatgtccc	atgtaaaacc	tgctacaaat	ggcagcttat	acatagcaat	ggtaaaatca	2340
tcctctggat	ttaggaattg	ctcttgtcat	acccccaagt	ttctaagatt	taagattctc	2400
cttactacta	tcctacgttt	aaatatcttt	gaaagtttgt	attaaatgtg	aattttaaga	2460
aataatattt	atattttctgt	aaatgtaaac	tgtgaagata	gttataaact	gaagcagata	2520
cctggaacca	cctaaagaac	ttccattttat	ggaggatttt	tttgccccctt	gtgtttggaa	2580
ttataaaata	taggtaaaag	tacgtaatta	aataatgttt	ttggtaaaaa	aaaaaaaaaa	2640
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aa	2692

<210> 130

<211> 2730

<212> DNA

<213> NM\_152872.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 3, mRNA

<400> 130  
cctacccgcg cgcaggccaa gttgctgaat caatggagcc ctccccaacc cgggcgttcc 60  
ccagcgaggc ttccttccca tcctcctgac caccggggct tttcgtgagc tcgtctctga 120  
tctcgcgcaa gagtgacaca caggtgttca aagacgcttc tggggagtga ggaagcggt 180  
ttacgagtga cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggc 240  
cagccctggc tgcccaggcg gagctgcctc ttctcccgcg ggttggtgga cccgctcagt 300  
acggagttgg ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg 360  
gacctccta cctctggttc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc 420  
ccaagtgact gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga 480  
gactcagaac ttggaaggcc tgcattcatga tggccaattc tgccataagc cctgtcctcc 540  
aggtgaaagg aaagctaggg actgcacagt caatggggat gaaccagact gcgtgccctg 600  
ccaagaaggg aaggagtaca cagacaaagc ccatttttct tccaaatgca gaagatgtag 660  
attgtgtgat gaaggacatg gcttagaagt ggaataaac tgcacccgga cccagaatac 720  
caagtgcaga tgtaaacc aaactttttt taactctact gtatgtgaac actgtgaccc 780  
ttgcaccaa tgtgaacatg gaatcatcaa ggaatgcaca ctaccagca acaccaagtg 840  
caaagaggaa ggatccagat ctaacttggg gtggctttgt cttcttcttt tgccaattcc 900  
actaattgtt tgggtgaaga gaaaggaagt acagaaaaca tgcagaaagc acagaaagga 960  
aaaccaaggt tctcatgaat ctccaacctt aaatcctatg ttgacttgag taaatatatc 1020  
accactattg ctggagtcac gacactaagt caagttaaag gctttgttcg aaagaatggt 1080  
gtcaatgaag ccaaaataga tgagatcaag aatgacaatg tccaagacac agcagaacag 1140  
aaagttcaac tgcttcgtaa ttggcatcaa cttcatggaa agaaagaagc gtatgacaca 1200  
ttgattaaag atctcaaaaa agccaatctt tgtactcttg cagagaaaat tcagactatc 1260  
atcctcaagg acattactag tgactcagaa aattcaaact tcagaaatga aatccaaagc 1320  
ttggtctaga gtgaaaaaca acaaattcag ttctgagtat atgcaattag tgtttgaaaa 1380  
gattcttaat agctggctgt aaatactgct tggtttttta ctgggtacat tttatcattt 1440  
attagcgtg aagagccaac atatttgtag atttttaata tctcatgatt ctgcctccaa 1500  
ggatgtttaa aatctagttg ggaaaacaaa cttcatcaag agtaaatgca gtggcatgct 1560  
aagtacccaa ataggagtgt atgcagagga tgaaagatta agattatgct ctggcatcta 1620  
acatatgatt ctgtagtatg aatgtaatca gtgtatgtta gtacaaatgt ctatccacag 1680  
gctaacccca ctctatgaat caatagaaga agctatgacc ttttgctgaa atatcagtta 1740  
ctgaacaggc aggccacttt gcctctaaat tacctctgat aattctagag attttaccat 1800

atttcta	aac	tttgttt	tata	actctg	agaa	gatcat	atattt	atgtaa	agta	tatgtatt	ttg	1860
agtgc	agaat	ttaaata	aagg	ctctac	ctca	aagac	ctttg	cacagt	tttat	tggtgt	cata	1920
ttata	caata	tttcaatt	gt	gaattc	acat	agaaa	acatt	aaattata	aat	gtttg	actat	1980
tatat	atgtg	tatgcatt	ttt	actgg	ctca	aactac	ctac	ttcttt	ctca	ggcat	caaaa	2040
gcatt	ttgag	caggag	agta	ttacta	gagc	tttgcc	acct	ctccatt	ttt	gcctt	gggtgc	2100
tcatt	ctta	at	ggccta	atgc	accccc	aaac	atggaa	atat	caccaaaa	tactta	atag	2160
tccac	caaaa	ggcaag	actg	ccctta	gaaa	ttctag	cctg	gtttgg	agat	acta	actgct	2220
ctcag	agaaa	gtagct	ttgt	gacat	gtcat	gaaccc	atgt	ttgca	atcaa	agat	gataaa	2280
atag	attctt	at	ttttt	cccc	caccccc	gaa	aatgtt	caat	aatgt	cccat	gtaaa	2340
ctaca	aatgg	cagctt	atac	atagca	atgg	taaa	atcatc	atctgg	attt	agga	attgct	2400
cttgt	catac	cccc	aagttt	ctaag	attta	agatt	ctct	tactac	tatc	ctacg	tttaa	2460
atat	ctttg	aagttt	gtat	taaat	gtgaa	tttta	agaaa	taatatt	ttat	attt	ctgtaa	2520
atgt	aaactg	tgaag	atagt	tataa	actga	agcag	atacc	tgga	accacc	taa	agaactt	2580
ccatt	tatgg	aggatt	ttttt	tgc	ccctt	gt	gtttg	gaatt	ataaa	atata	ggtaa	2640
cgta	attaaa	taat	gttttt	ggt	aaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	2700
aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	aaaaaaaa	2730

<210> 131

<211> 2563

<212> DNA

<213> NM\_152874.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 8, mRNA

<400>	131	cctaccc	gcg	cg	cagg	ccaa	gtt	gct	gaat	caat	ggag	cc	ctcccc	aacc	cg	ggc	gttcc	60
ccagc	gag	gc	ttc	ctt	cccc	tcct	cct	gac	cacc	ggg	gct	ttc	g	tg	gag	c	tcgt	120
tctcg	c	gaa	gag	t	gac	aca	cagg	t	gtt	ca	aag	ac	gtt	c	t	g	ggg	180
ttac	g	agt	ga	ctt	gg	ct	g	g	g	g	g	g	cac	g	g	a	acc	240
cag	cc	t	g	g	c	g	g	g	g	g	g	g	g	g	g	g	g	300
ac	g	g	g	g	g	g	g	g	g	g	g	g	g	g	g	g	g	360
g	ac	cc	t	c	ct	a	g	g	g	g	g	g	g	g	g	g	g	420
cca	g	t	g	a	g	t	g	a	g	g	g	g	g	g	g	g	g	480
g	a	c	t	c	a	a	c	a	a	c	a	a	c	a	a	c	a	540
agg	t	g	a	a	a	g	g	g	g	g	g	g	g	g	g	g	g	600
cca	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	660



attgtgtgat gaaggacatg atgtgaacat ggaatcatca aggaatgcac actcaccagc	720
aacaccaagt gcaaagagga aggatccaga tctaacttgg ggtggctttg tcttcttctt	780
ttgccaatc cactaattgt ttgggggaaac agtggcaata aatttatctg atgttgactt	840
gagtaaatat atcaccacta ttgctggagt catgacacta agtcaagtta aaggctttgt	900
tcgaaagaat ggtgtcaatg aagccaaaat agatgagatc aagaatgaca atgtccaaga	960
cacagcagaa cagaaagttc aactgcttcg taattggcat caacttcatg gaaagaaaga	1020
agcgtatgac acattgatta aagatctcaa aaaagccaat ctttgactc ttgcagagaa	1080
aattcagact atcatcctca aggacattac tagtgactca gaaaattcaa acttcagaaa	1140
tgaaatccaa agcttggctt agagtgaaaa acaacaaatt cagttctgag tatatgcaat	1200
tagtgtttga aaagattctt aatagctggc tgtaaatact gcttggtttt ttactgggta	1260
cattttatca tttattagcg ctgaagagcc aacatatttg tagattttta atatctcatg	1320
attctgcctc caaggatggt taaaatctag ttgggaaaac aaacttcac aagagtaaat	1380
gcagtggcat gctaagtacc caaataggag tgtatgcaga ggatgaaaga ttaagattat	1440
gctctggcat ctaacatatg attctgtagt atgaatgtaa tcagtgtatg ttagtacaaa	1500
tgtctatcca caggctaacc ccactctatg aatcaataga agaagctatg accttttgct	1560
gaaatatcag ttactgaaca ggcaggccac ttgacctcta aattacctct gataattcta	1620
gagattttac catatttcta aactttgttt ataactctga gaagatcata tttatgtaaa	1680
gtatatgtat ttgagtgcag aattttaata aggtctctacc tcaaagacct ttgcacagtt	1740
tattgggtgtc atattatata atattticaat tgtgaattca catagaaaac attaaattat	1800
aatgtttgac tattatatat gtgtatgcat ttacttggct caaaactacc tacttctttc	1860
tcaggcatca aaagcatttt gagcaggaga gtattactag agctttgcc cctctccatt	1920
tttgcccttg tgctcatctt aatggcctaa tgcaccccca aacatggaaa tatcaccaaa	1980
aaatacttaa tagtccacca aaaggcaaga ctgcccttag aaattctagc ctggtttgga	2040
gatactaact gctctcagag aaagtagctt tgtgacatgt catgaacca tgtttgcaat	2100
caaagatgat aaaatagatt cttatttttc cccaccccc gaaaatgttc aataatgtcc	2160
catgtaaaac ctgctacaaa tggcagctta tacatagcaa tggtaaaatc atcatctgga	2220
tttaggaatt gctcttgta taccaccaag ttcttaagat ttaagattct ccttactact	2280
atcctacgtt taaatatctt tgaaagtttg tattaaatgt gaattttaag aaataatatt	2340
tatatttctg taaatgtaaa ctgtgaagat agttataaac tgaagcagat acctggaacc	2400
acctaaagaa cttccattta tggaggattt ttttgcccct tgtgtttgga attataaaat	2460
ataggtaaaa gtacgtaatt aaataatgtt ttgggtaaaa aaaaaaaaaa aaaaaaaaaa	2520
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa	2563

<210> 132

<211> 2445

<212> DNA

<213> NM\_152876.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 6, mRNA

<400> 132

```
cctacccgcg cgcaggccaa gttgctgaat caatggagcc ctccccaacc cgggcgttcc 60
ccagcgagggc ttccttccca tctcctgac caccggggct tttcgtgagc tcgtctctga 120
tctcgcgcaa gagtgcacac caggtgttca aagacgcttc tggggagtga ggggaagcgg 180
ttacgagtga cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggg 240
cagccctggc tgcccaggcg gagctgcctc ttctcccgcg ggttggtgga cccgctcagt 300
acggagtggg ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg 360
gacctccta cctctggttc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc 420
ccaagtgact gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga 480
gactcagaac ttggaaggcc tgcattcatg tggccaattc tgccataagc cctgtcctcc 540
agatgtgaac atggaatcat caaggaatgc acactcacca gcaacaccaa gtgcaaagag 600
gaagtgaaga gaaaggaagt acagaaaaca tgcagaaagc acagaaagga aaaccaaggt 660
tctcatgaat ctccaacctt aaatcctgaa acagtggcaa taaatttatc tgatgttgac 720
ttgagtaa atataccac tattgctgga gtcattgacac taagtcaagt taaaggcttt 780
gttcgaaaga atggtgtcaa tgaagccaaa atagatgaga tcaagaatga caatgtccaa 840
gacacagcag aacagaaagt tcaactgctt cgtaattggc atcaacttca tggaaagaaa 900
gaagcgtatg acacattgat taaagatctc aaaaaagcca atctttgtac tcttgagag 960
aaaattcaga ctatcatcct caaggacatt actagtgact cagaaaattc aaacttcaga 1020
aatgaaatcc aaagcttggg ctagagtga aaacaacaaa ttcagttctg agtatatgca 1080
attagtgttt gaaaagattc ttaatagctg gctgtaaata ctgcttgggt ttttactggg 1140
tacattttat catttattag cgctgaagag ccaacatatt tgtagatttt taatatctca 1200
tgattctgcc tccaaggatg tttaaaatct agttgggaaa acaaacttca tcaagagtaa 1260
atgcagtggc atgctaagta cccaaatagg agtgtatgca gaggatgaaa gattaagatt 1320
atgctctggc atctaacata tgattctgta gtatgaatgt aatcagtgta tgtagtaca 1380
aatgtctatc cacaggctaa cccactcta tgaatcaata gaagaagcta tgaccttttg 1440
ctgaaatatc agttactgaa caggcaggcc actttgcctc taaattacct ctgataattc 1500
tagagatttt accatatttc taaactttgt ttataactct gagaagatca tatttatgta 1560
aagtatatgt atttgagtgc agaatttaaa taaggctcta cctcaaagac ctttgcacag 1620
tttattgggtg tcatattata caatatttca attgtgaatt cacatagaaa acattaaatt 1680
```

ataatgtttg actattatat atgtgtatgc attttactgg ctcaaaaacta cctacttctt	1740
tctcaggcat caaaagcatt ttgagcagga gagtattact agagctttgc cacctctcca	1800
tttttgccct ggtgctcatc ttaatggcct aatgcacccc caaacatgga aatatcacca	1860
aaaaatactt aatagtccac caaaaggcaa gactgccctt agaaattcta gcctggtttg	1920
gagatactaa ctgctctcag agaaagtagc tttgtgacat gtcatgaacc catgtttgca	1980
atcaaagatg ataaaataga ttcttatttt tccccaccc ccgaaaatgt tcaataatgt	2040
cccatgtaaa acctgctaca aatggcagct tatacatagc aatggtaaaa tcatcatctg	2100
gatttaggaa ttgctcttgt catacccca agtttctaag atttaagatt ctccttacta	2160
ctatcctacg tttaaatatac tttgaaagtt tgtattaaat gtgaatttta agaaataata	2220
tttatatttc tgtaaagtga aactgtgaag atagttataa actgaagcag atacctggaa	2280
ccacctaaag aacttccatt tatggaggat tttttgccc cttgtgtttg gaattataaa	2340
atataggtaa aagtacgtaa ttaaataatg tttttggtaa aaaaaaaaaa aaaaaaaaaa	2400
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa	2445

<210> 133

<211> 2508

<212> DNA

<213> NM\_152877.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 7, mRNA

<400> 133	
cctacccgcg cgcaggccaa gttgctgaat caatggagcc ctccccaacc cgggcgttcc	60
ccagcgaggc ttctttccca tcctcctgac caccggggct tttcgtgagc tcgtctctga	120
tctcgcgcaa gagtgacaca cagggtgttca aagacgcttc tggggagtga gggaagcgg	180
ttacgagtga cttggctgga gcctcagggg cgggcactgg cacggaacac accctgaggc	240
cagccctggc tgcccaggcg gagctgcctc ttctcccgcg ggttgggtgga cccgctcagt	300
acggagttgg ggaagctctt tcacttcgga ggattgctca acaaccatgc tgggcatctg	360
gaccctccta cctctggttc ttacgtctgt tgctagatta tcgtccaaaa gtgttaatgc	420
ccaagtgact gacatcaact ccaagggatt ggaattgagg aagactgtta ctacagttga	480
gactcagaac ttggaaggcc tgcacatga tggccaattc tgccataagc cctgtcctcc	540
agatgtgaac atggaatcat caaggaatgc aactcacca gcaacaccaa gtgcaaagag	600
gaaggatcca gatctaactt ggggtggcct tgtcttcttc ttttgccaat tccactaatt	660
gtttgggtga agagaaagga agtacagaaa acatgcagaa agcacagaaa ggaaaaccaa	720
ggttctcatg aatctccaac cttaaatcct gaaacagtgg caataaattt atctgatgtt	780
gacttgagta aatatatcac cactattgct ggagtcatga cactaagtca agttaaggc	840

tttgttcgaa agaatggtgt caatgaagcc aaaatagatg agatcaagaa tgacaatgtc	900
caagacacag cagaacagaa agttcaactg cttcgtaatt ggcacact tcatggaaag	960
aaagaagcgt atgacacatt gattaaagat ctcaaaaaag ccaatctttg tactcttgca	1020
gagaaaattc agactatcat cctcaaggac attactagtg actcagaaaa ttcaaacttc	1080
agaaatgaaa tccaaagctt ggtctagagt gaaaaacaac aaattcagtt ctgagtatat	1140
gcaattagtg tttgaaaaga ttcttaatag ctggctgtaa atactgcttg gttttttact	1200
gggtacattt tatcatttat tagcgctgaa gagccaacat atttgtagat ttttaatatc	1260
tcatgattct gcctccaagg atgttttaaaa tctagttggg aaaacaaact tcatcaagag	1320
taaatgcagt ggcacgctaa gtacccaaat aggagtgtat gcagaggatg aaagattaag	1380
attatgctct ggcacctaac atatgattct gtagtatgaa tgtaatcagt gtatgttagt	1440
acaaatgtct atccacaggc taacccact ctatgaatca atagaagaag ctatgacctt	1500
ttgctgaaat atcagttact gaacaggcag gccactttgc ctctaaatta cctctgataa	1560
ttctagagat ttaccatat ttctaaactt tgtttataac tctgagaaga tcatatttat	1620
gtaaagtata tgtatttgag tgcagaattt aaataaggct ctacctcaa gacctttgca	1680
cagtttattg gtgtcatatt atacaatatt tcaattgtga attcacatag aaacattaa	1740
attataatgt ttgactatta tatatgtgta tgcattttac tggctcaaaa ctacctactt	1800
ctttctcagg catcaaaagc attttgagca ggagagtatt actagagctt tgccacctct	1860
ccatttttgc cttgggtgctc atcttaatgg cctaatgcac ccccaaacat ggaaatatca	1920
ccaaaaaata cttaatagtc caccaaaagg caagactgcc cttagaaatt ctagcctggt	1980
ttggagatac taactgctct cagagaaagt agctttgtga catgtcatga acccatgttt	2040
gcaatcaaag atgataaaat agattcttat ttttcccca ccccgaaaa tgttcaataa	2100
tgtcccatgt aaaacctgct acaaatggca gcttatacat agcaatggta aaatcatcat	2160
ctggatttag gaattgctct tgtcataccc ccaagtttct aagatttaag attctcctta	2220
ctactatcct acgtttaaat atctttgaaa gtttgtatta aatgtgaatt ttaagaaata	2280
atatttatat ttctgtaaat gtaaactgtg aagatagtta taaactgaag cagataacctg	2340
gaaccaccta aagaacttcc atttatggag gatttttttg ccccttgtgt ttggaattat	2400
aaaatatagg taaaagtacg taattaaata atgttttttg taaaaaaaaa aaaaaaaaaa	2460
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa	2508

<210> 134

<211> 2583

<212> DNA

<213> NM\_152875.1| Homo sapiens tumor necrosis factor receptor superfamily, member 6 (TNFRSF6), transcript variant 5, mRNA

<400> 134

cctacccgcg	cgagggccaa	gttgctgaat	caatggagcc	ctccccaacc	cgggcgttcc	60
ccagcgaggc	ttccttccca	tcctcctgac	caccggggct	tttcgtgagc	tcgtctctga	120
tctcgcgcaa	gagtgcacac	caggtgttca	aagacgcttc	tggggagtga	gggaagcggc	180
ttacgagtga	cttggctgga	gcctcagggg	cgggcactgg	cacggaacac	accctgaggc	240
cagccctggc	tgcccaggcg	gagctgcctc	ttctcccgcg	ggttgggtgga	cccgtcagc	300
acggagttgg	ggaagctctt	tcacttcgga	ggattgctca	acaaccatgc	tgggcactctg	360
gaccctccta	cctctgggtc	ttacgtctgt	tgctagatta	tcgtccaaaa	gtgttaatgc	420
ccaagtgcact	gacatcaact	ccaagggatt	ggaattgagg	aagactgtta	ctacagttga	480
gactcagaac	ttggaaggcc	tgcatcatga	tggccaattc	tgccataagc	cctgtcctcc	540
aggtgaaagg	aaagctaggg	actgcacagt	caatggggat	gaaccagact	gcgtgccctg	600
ccaagaaggg	aaggagtaca	cagacaaagc	ccatTTTTtct	tccaaatgca	gaagatgtag	660
attgtgtgat	gaaggacatg	atgtgaacat	ggaatcatca	aggaatgcac	actcaccagc	720
aacaccaagt	gcaaagagga	agtgaagaga	aaggaagtac	agaaaacatg	cagaaagcac	780
agaaaggaaa	accaaggttc	tcatgaatct	ccaaccttaa	atcctgaaac	agtggcaata	840
aatttatctg	atgttgactt	gagtaaatat	atcaccacta	ttgctggagt	catgacacta	900
agtcaagtta	aaggctttgt	tcgaaagaat	ggtgtcaatg	aagccaaaat	agatgagatc	960
aagaatgaca	atgtccaaga	cacagcagaa	cagaaagtcc	aactgcttcg	taattggcat	1020
caacttcatg	gaaagaaaga	agcgtatgac	acattgatta	aagatctcaa	aaaagccaat	1080
ctttgtactc	ttgcagagaa	aattcagact	atcatcctca	aggacattac	tagtgactca	1140
gaaaattcaa	acttcagaaa	tgaaatccaa	agcttgggtc	agagtgaaaa	acaacaaatt	1200
cagttctgag	tatatgcaat	tagtgtttga	aaagattcct	aatagctggc	tgtaaatact	1260
gcttggtttt	ttactgggta	cattttatca	tttattagcg	ctgaagagcc	aacatatttg	1320
tagattttta	atatctcatg	attctgcctc	caaggatggt	taaaatctag	ttgggaaaac	1380
aaacttcatc	aagagtaaat	gcagtggcat	gctaagtacc	caaataggag	tgtatgcaga	1440
ggatgaaaga	ttaagattat	gctctggcat	ctaacatatg	attctgtagt	atgaatgtaa	1500
tcagtgtatg	ttagtacaaa	tgtctatcca	caggctaacc	ccactctatg	aatcaataga	1560
agaagctatg	accttttgct	gaaatatcag	ttactgaaca	ggcaggccac	tttgccctca	1620
aattacctct	gataattcta	gagattttac	catatttcta	aactttgttt	ataactctga	1680
gaagatcata	tttatgtaaa	gtatatgtat	ttgagtgcag	aatttaaata	aggctctacc	1740
tcaaagacct	ttgcacagtt	tattgggtgc	atattataca	atatttcaat	tgtgaattca	1800
catagaaaac	attaaattat	aatgtttgac	tattatatat	gtgtatgcat	tttactggct	1860
caaaactacc	tacttctttc	tcaggcatca	aaagcatttt	gagcaggaga	gtattactag	1920

```

agctttgccca cctctccatt tttgccttgg tgctcatctt aatggcctaa tgcaccccca 1980
aacatggaaa tatcaccaaa aaatacttaa tagtccacca aaaggcaaga ctgcccttag 2040
aaattctagc ctggtttgga gatactaact gctctcagag aaagtagctt tgtgacatgt 2100
catgaaccca tgtttgcaat caaagatgat aaaatagatt cttatttttc cccaccccc 2160
gaaaatgttc aataatgtcc catgtaaaac ctgctacaaa tggcagctta tacatagcaa 2220
tggtaaaatc atcatctgga tttaggaatt gctcttgta taccaccaag tttctaagat 2280
ttaagattct ctttactact atcctacgtt taaatatctt tgaaagtttg tattaaatgt 2340
gaattttaag aaataatatt tatatttctg taaatgtaaa ctgtgaagat agttataaac 2400
tgaagcagat acctggaacc acctaaagaa cttccattta tggaggattt ttttgcccct 2460
tgtgtttgga attataaaat ataggtaaaa gtacgtaatt aaataatgtt tttggtaaaa 2520
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2580
aaa 2583

```

<210> 135

<211> 316

<212> DNA

<213> >gi|13310411|gb|AF333388.1|AF333388 Homo sapiens metallothionein 1H-like protein mRNA, complete cds

```

<400> 135
cctcttctct tctcgcttgg gaacgccggt ctcacctcgg cttgcaatgg accccaactg 60
ctcctgcgcc gctggaggct cctacgcctg cgccggctcc tgcaagtgca aaaagtgcaa 120
atgcacctcc tgcaagaaga gctgctgctc ctgttgcccc ctgggctgtg ccaagtgtgc 180
ccagggctgc atccgcaaag gggcttcgga aaagtgcagc tgctgtgcct gatgtcggga 240
ctgccctgct ctcggatgaa aacagaatga cacgtaaagt ccgggatttt tttttctaca 300
actccgactc atttgc 316

```

<210> 136

<211> 3145

<212> DNA

<213> NM\_000251. Homo sapiens muts...[gi:4557760]

```

<400> 136
ggcgggaaac agcttagtgg gtgtggggtc gcgcattttc ttcaaccagg aggtgaggag 60
gtttcgacat ggcggtgcag ccgaaggaga cgctgcagtt ggagagcgcg gccgaggtcg 120
gcttcgtgcg cttcttttcag ggcatgccgg agaagccgac caccacagtg cgctttttcg 180

```

accggggcga	cttctatacg	gcgcacggcg	aggacgcgct	gctggccgcc	cgggaggtgt	240
tcaagaccca	gggggtgatc	aagtacatgg	ggccggcagg	agcaaagaat	ctgcagagtg	300
ttgtgcttag	taaaatgaat	tttgaatctt	ttgtaaaaga	tcttcttctg	gttcgtcagt	360
atagagttga	agttttataag	aatagagctg	gaaataaggc	atccaaggag	aatgattggt	420
at ttggcata	taaggcttct	cctggcaatc	tctctcagtt	tgaagacatt	ctctttggta	480
acaatgatat	gtcagcttcc	attggtgttg	tgggtgttaa	aatgtccgca	gttgatggcc	540
agagacaggt	tggagttggg	tatgtggatt	ccatacagag	gaaactagga	ctgtgtgaat	600
tccctgataa	tgatcagttc	tccaatcttg	aggctctcct	catccagatt	ggaccaaagg	660
aatgtgtttt	acccggagga	gagactgctg	gagacatggg	gaaactgaga	cagataattc	720
aaagaggagg	aattctgatc	acagaaagaa	aaaaagctga	cttttccaca	aaagacattt	780
atcaggacct	caaccggttg	ttgaaaggca	aaaagggaga	gcagatgaat	agtgtgttat	840
tgccagaaat	ggagaatcag	gttgagttt	catcactgtc	tgcggtaatc	aagtttttag	900
aactcttatc	agatgattcc	aactttggac	agtttgaact	gactactttt	gacttcagcc	960
agtatatgaa	attggatatt	gcagcagtc	gagcccttaa	cctttttcag	ggttctgttg	1020
aagataccac	tggctctcag	tctctggctg	ccttgctgaa	taagtgtaaa	acccctcaag	1080
gacaaagact	tgtaaccag	tggattaagc	agcctctcat	ggataagaac	agaatagagg	1140
agagattgaa	tttagtggaa	gcttttgtag	aagatgcaga	attgaggcag	actttacaag	1200
aagatttact	tcgtcgattc	ccagatctta	accgacttgc	caagaagttt	caaagacaag	1260
cagcaaactt	acaagattgt	taccgactct	atcagggtat	aatcaacta	cctaattgta	1320
tacaggctct	ggaaaaacat	gaaggaaaac	accagaaatt	attgttggca	gtttttgtga	1380
ctcctcttac	tgatcttcgt	tctgacttct	ccaagtttca	ggaaatgata	gaaacaactt	1440
tagatatgga	tcagggtggaa	aaccatgaat	tccttgtaaa	accttcattt	gatcctaatt	1500
tcagtgaatt	aagagaaata	atgaatgact	tggaaaagaa	gatgcagtc	acattaataa	1560
gtgcagccag	agatcttggc	ttggaccctg	gcaaacagat	taaactggat	tccagtgcac	1620
agttttggata	ttactttcgt	gtaacctgta	aggaagaaaa	agtccttcgt	aacaataaaa	1680
acttttagtac	tgtagatatc	cagaagaatg	gtgttaaatt	taccaacagc	aaattgactt	1740
ctttaaatga	agagtatacc	aaaaataaaa	cagaatatga	agaagcccag	gatgccattg	1800
ttaaagaaat	tgtcaatatt	tcttcaggct	atgtagaacc	aatgcagaca	ctcaatgatg	1860
tgttagctca	gctagatgct	gttgtcagct	ttgctcacgt	gtcaaattgga	gcacctgttc	1920
catatgtacg	accagccatt	ttggagaaag	gacaaggaag	aattatatta	aaagcatcca	1980
ggcatgcttg	tgttgaagtt	caagatgaaa	ttgcatttat	tcctaattgac	gtatactttg	2040
aaaaagataa	acagatgttc	cacatcatta	ctggccccaa	tatgggaggt	aatcaacat	2100
atattcgaca	aactgggggtg	atagtactca	tggcccaaat	tgggtgtttt	gtgccatgtg	2160
agtcagcaga	agtggtccatt	gtggactgca	tcttagcccg	agtaggggct	ggtgacagtc	2220

aattgaaagg agtctccacg ttcattggctg aaatgttggg aactgcttct atcctcaggt	2280
ctgcaaccaa agattcatta ataatcatag atgaattggg aagaggaact tctacctacg	2340
atggatttgg gtttagcatgg gctatatcag aatacattgc aacaaagatt ggtgcttttt	2400
gcatgtttgc aacccatttt catgaactta ctgccttggc caatcagata ccaactgtta	2460
ataatctaca tgtcacagca ctccacctg aagagacctt aactatgctt tatcaggtga	2520
agaaagggtg ctgtgatcaa agttttggga ttcatgttgc agagcttgct aatttcctta	2580
agcatgtaat agagtgtgct aaacagaaag ccctggaact tgaggagttt cagtatatgt	2640
gagaatcgca aggatatgat atcatggaac cagcagcaaa gaagtgctat ctggaaagag	2700
agcaagggtga aaaaattatt caggagtcc tgtccaagggt gaaacaaatg ccctttactg	2760
aaatgtcaga agaaaacatc acaataaagt taaaacagct aaaagctgaa gtaatagcaa	2820
agaataatag ctttgtaaat gaaatcattt cacgaataaa agttactacg tgaaaaatcc	2880
cagtaatgga atgaaggtaa tattgataag ctattgtctg taatagtttt atattgtttt	2940
atattaaccc tttttccata gtgttaactg tcagtgccca tgggctatca acttaataag	3000
atatttagta atattttact ttgaggacat tttcaaagat ttttattttg aaaaatgaga	3060
gctgtaactg aggactgttt gcaattgaca taggcaataa taagtgatgt gctgaatttt	3120
ataaataaaa tcatgtagtt tgtgg	3145

<210> 137

<211> 3239

<212> DNA

<213> NM\_000534. Homo sapiens PMS1...[gi:53729349]

<400> 137	
ctcgctgccg gcggattggc tgcgagcagc gccaatctca cgttgccccg gggcgaggcg	60
ggactcagtg ccgcgctctc tgcacccgct ctgccgcgcg cgtgcgtgct ggggtgcgggt	120
gcgggtgctg ggttgggcct gcgcacgctg tgagacgctg gctgcttgct gctagtggat	180
ggtaattgcc tgcctcgctg tagcaggaag ctgctctgtt aaaagcgaaa atgaaacaat	240
tgcctgcggc aacagttcga ctcccttcaa gttctcagat catcacttcg gtggtcagt	300
ttgtaaaaga gcttattgaa aactccttgg atgctggtgc cacaagcgta gatgttaa	360
tggagaacta tggatttgat aaaattgagg tgcgagataa cggggagggt atcaaggctg	420
ttgatgcacc tgtaattggc atgaagtact acacctcaaa aataaatagt catgaagatc	480
ttgaaaattt gacaacttac ggttttcgtg gagaagcctt ggggtcaatt tgttgatat	540
ctgagggttt aattacaaca agaacggctg ctgataattt tagcaccag tatgttttag	600
atggcagtg ccacatactt tctcagaaac cttcacatct tgggtcaagggt acaactgtaa	660



ctgctttaag attattttaag aatctacctg taagaaagca gttttactca actgcaaaaa	720
aatgtaaaga tgaaataaaaa aagatccaag atctcctcat gagctttggt atccttaaac	780
ctgacttaag gattgtcttt gtacataaca aggagttat ttggcagaaa agcagagtat	840
cagatcacia gatggctctc atgtcagttc tggggactgc tggtatgaac aatatggaat	900
cctttcagta ccactctgaa gaatctcaga tttatctcag tggatttctt ccaaagtgtg	960
atgcagacca ctcttttact agtctttcaa caccagaaag aagtttcatc ttcataaaca	1020
gtcgaccagt acatcaaaaa gatattctta agttaatccg acatcattac aatctgaaat	1080
gcctaaagga atctactcgt ttgtatcctg ttttctttct gaaaatcgat gttcctacag	1140
ctgatgttga tgtaaattta acaccagata aaagccaagt attattacaa aataaggaat	1200
ctgttttaat tgctcttgaa aatctgatga cgacttgta tggaccatta cctagtacaa	1260
attcttatga aaataataaa acagatgttt ccgcagctga catcgttctt agtaaacag	1320
cagaaacaga tgtgcttttt aataaagtgg aatcatctgg aaagaattat tcaaagtgtg	1380
atacttcagt cattccattc caaaatgata tgcataatga tgaatctgga aaaaacactg	1440
atgattgttt aaatcaccag ataagtattg gtgactttgg ttatgggtcat tgtagtagtg	1500
aaatttctaa cattgataaa aacactaaga atgcatttca ggacatttca atgagtaatg	1560
tatcatggga gaactctcag acggaatata gtaaaacttg ttttataagt tccgttaagc	1620
acaccagtc agaaaatggc aataaagacc atatagatga gagtggggaa aatgaggaag	1680
aagcaggtct tgaaaactct tcggaaattt ctgcagatga gtggagcagg ggaaatatac	1740
ttaaaaattc agtgggagag aatattgaac ctgtgaaaat tttagtgcct gaaaaaagtt	1800
taccatgtaa agtaagtaat aataattatc caatccctga acaaataaat cttaatgaag	1860
attcatgtaa caaaaaatca aatgtaatag ataataaatc tggaaaagtt acagcttatg	1920
atttacttag caatcgagta atcaagaaac ccatgtcagc aagtgtcttt tttgttcaag	1980
atcatcgtcc tcagtttctc atagaaaatc ctaagactag tttagaggat gcaacactac	2040
aaattgaaga actgtggaag acattgagtg aagaggaaaa actgaaatat gaagagaagg	2100
ctactaaaga cttggaacga tacaatagtc aaatgaagag agccattgaa caggagtcac	2160
aaatgtcact aaaagatggc agaaaaaaga taaaaccac cagcgcatgg aatttgcccc	2220
agaagcacia gttaaaaacc tcattatcta atcaacaaaa acttgatgaa ctccttcagt	2280
cccaaattga aaaaagaagg agtcaaaata ttaaatggt acagatcccc ttttctatga	2340
aaaacttaaa aataaatttt aagaacaaa acaaagttga cttagaagag aaggatgaac	2400
cttgcttgat ccacaatctc aggtttcctg atgcatggct aatgacatcc aaaacagagg	2460
taatgttatt aaatccatat agagtagaag aagccctgct atttaaaaga cttcttgaga	2520
atcataaact tcctgcagag ccactggaaa agccaattat gttaacagag agtcttttta	2580
atggatctca ttatttagac gttttatata aaatgacagc agatgaccaa agatacagtg	2640
gatcaactta cctgtctgat cctcgtctta cagcgaatgg tttcaagata aaattgatac	2700

caggagtttc aattactgaa aattacttgg aaatagaagg aatggctaatt tgtctcccat	2760
tctatggagt agcagattta aaagaaattc ttaatgctat attaaacaga aatgcaaagg	2820
aagtttatga atgtagacct cgcaaagtga taagttatth agagggagaa gcagtgcgtc	2880
tatccagaca attacccatg tacttatcaa aagaggacat ccaagacatt atctacagaa	2940
tgaagcacca gtttggaaat gaaattaaag agtgtgttca tggtcgcca ttttttcac	3000
atttaaccta tcttcagaa actacatgat taaatatgtt taagaagatt agttaccatt	3060
gaaattgggt ctgtcataaa acagcatgag tctggtttta aattatcttt gtattatgtg	3120
tcacatgggt atttttttaa tgaggattca ctgacttgtt tttatattga aaaaagttcc	3180
acgtattgta gaaaacgtaa ataaactaat atagactatt caaaaaaaaa aaaaaaaaa	3239

<210> 138

<211> 2771

<212> DNA

<213> NM\_000535. Homo sapiens PMS2...[gi:11125773]

<400> 138	
cgaggcggat cgggtgttgc atccatggag cgagctgaga gctcgagtac agaacctgct	60
aaggccatca aacctattga tcggaagtca gtccatcaga tttgctctgg gcaggtggt	120
ctgagtctaa gactgcggt aaaggagtta gtagaaaaca gtctggatgc tggtgccact	180
aatattgatc taaagcttaa ggactatgga gtggatctta ttgaagtttc agacaatgga	240
tgtggggtag aagaagaaaa cttcgaaggc ttaactctga aacatcacac atctaagatt	300
caagagtttg ccgacctaac tcagggtgaa acttttggct ttcgggggga agctctgagc	360
tcactttgtg cactgagcga tgtcaccatt tctacctgcc acgcatcggc gaaggttgga	420
actcgactga tgtttgatca caatgggaaa attatccaga aaaccccta ccccccccc	480
agagggacca cagtcagcgt gcagcagtta tttccacac tacctgtgcg ccataaggaa	540
tttcaaagga atattaagaa ggagtatgcc aaaatgggtcc aggtcttaca tgcatactgt	600
atcatttcag caggcatccg tgtaagttgc accaatcagc ttggacaagg aaaacgacag	660
cctgtggtat gcacagggtg aagccccagc ataaaggaaa atatcggctc tgtgtttggg	720
cagaagcagt tgcaaagcct cattcctttt gttcagctgc cccctagtga ctccgtgtgt	780
gaagagtacg gtttgagctg ttcggatgct ctgcataatc tttttacat ctccagtttc	840
atttcacaat gcacgcatgg agttggaagg agttcaacag acagacagtt tttctttatc	900
aaccggcggc cttgtgaccc agcaaaggtc tgcagactcg tgaatgaggt ctaccacatg	960
tataatcgac accagtatcc atttgttgtt cttaacattt ctgttgattc agaatgcgtt	1020
gatatcaatg tttactccaga taaaaggcaa attttgctac aagaggaaaa gcttttgttg	1080

gcagttttaa agacctcttt gataggaatg tttgatagtg atgtcaacaa gctaaatgtc	1140
agtcagcagc cactgctgga tgttgaaggt aacttaataa aaatgcatgc agcggatttg	1200
gaaaagccca tggtagaaaa gcaggatcaa tcccccttcat taaggactgg agaagaaaaa	1260
aaagacgtgt ccatttccag actgcgagag gcctttttctc ttcgtcacac aacagagaac	1320
aagcctcaca gccc aaagac tccagaacca agaaggagcc ctctaggaca gaaaaggggt	1380
atgctgtctt ctagcacttc aggtgccatc tctgacaaag gcgtcctgag acctcagaaa	1440
gaggcagtga gttccagtca cggacccagt gaccctacgg acagagcggg ggtggagaag	1500
gactcggggc acggcagcac ttccgtggat tctgaggggt tcagcatccc agacacgggc	1560
agtcactgca gcagcgagta tgcggccagc tccccagggg acaggggctc gcaggaacat	1620
gtggactctc aggagaaagc gcctgaaact gacgactctt tttcagatgt ggactgccat	1680
tcaaaccagg aagataccgg atgtaaatth cgagttttgc ctcagccaac taatctcgca	1740
accccaaaca caaagcgttt taaaaaagaa gaaattcttt ccagttctga catttgtcaa	1800
aagttagtaa atactcagga catgtcagcc tctcaggttg atgtagctgt gaaaattaat	1860
aagaaagttg tgccccctgga cttttctatg agttcttttag ctaaacgaat aaagcagtta	1920
catcatgaag cacagcaaag tgaaggggaa cagaattaca ggaagtttag ggcaaagatt	1980
tgtcctggag aaaatcaagc agccgaagat gaactaagaa aagagataag taaaacgatg	2040
tttgagaaa tggaaatcat tggtcagttt aacctgggat ttataataac caaactgaat	2100
gaggatatct tcatagtgga ccagcatgcc acggacgaga agtataactt cgagatgctg	2160
cagcagcaca ccgtgctcca ggggcagagg ctcatagcac ctcagactct caacttaact	2220
gctgttaatg aagctgttct gatagaaaat ctggaaatat ttagaaagaa tggctttgat	2280
tttgttatcg atgaaaatgc tccagtcact gaaagggcta aactgatttc cttgccaaact	2340
agtaaaaact ggaccttcgg accccaggac gtcgatgaac tgatcttcat gctgagcgac	2400
agccctgggg tcatgtgccg gccttcccg gtcaagcaga tgtttgctc cagagcctgc	2460
cggaagtcgg tgatgattgg gactgctctt aacacaagcg agatgaagaa actgatcacc	2520
cacatggggg agatggacca cccctggaac tgtcccatg gaaggccaac catgagacac	2580
atcgccaacc tgggtgtcat ttctcagaac tgaccgtagt cactgtatgg aataattgg	2640
tttatcgag atttttatgt tttgaaagac agagtcttca ctaacctttt ttgttttaaa	2700
atgaaacctg ctacttaaaa aaaatacaca tcacacccat ttaaaagtga tcttgagaac	2760
cttttcaaac c	2771

<210> 139

<211> 4264

<212> DNA

<213> NM\_000179. Homo sapiens mutS...[gi:4504190]

<400> 139						
atttccccgcc	agcaggagcc	gcgcggtaga	tgcggtgctt	ttaggagctc	cgtccgacag	60
aacggttggg	ccttgccggc	tgtcggtagt	tcgcgacaga	gcaccctgta	cagcttcttc	120
cccaagtctc	cggcgctgag	tgatgccaac	aaggcctcgg	ccagggcctc	acgcgaaggc	180
ggccgtgccg	ccgctgcccc	cggggcctct	ccttccccag	gcgggggatgc	ggcctggagc	240
gaggctgggc	ctgggcccag	gcccttggcg	cgatccgcgt	caccgcccac	ggcgaagaac	300
ctcaacggag	ggctgcggag	atcggtagcg	cctgtgcccc	ccaccagttg	tgacttctca	360
ccaggagatt	tggtttgggc	caagatggag	ggttacccct	ggtggccttg	tctggtttac	420
aaccaccctt	ttgatggaac	attcatccgc	gagaaagggg	aatcagtcgg	tgttcatgta	480
cagttttttg	atgacagccc	aacaaggggc	tgggttagca	aaaggctttt	aaagccatat	540
acaggttcaa	aatcaaagga	agcccagaag	ggaggtcatt	tttacagtgc	aaagcctgaa	600
atactgagag	caatgcaacg	tgcagatgaa	gccttaaata	aagacaagat	taagaggctt	660
gaattggcag	tttgtgatga	gccctcagag	ccagaagagg	aagaagagat	ggaggtaggc	720
acaacttacg	taacagataa	gagtgaagaa	gataatgaaa	ttgagagtga	agaggaagta	780
cagcctaaga	cacaaggatc	taggcgaagt	agccgcccac	taaaaaacg	aagggtcata	840
tcagattctg	agagtgcacat	tgggtggctct	gatgtggaat	ttaagccaga	cactaaggag	900
gaaggaagca	gtgatgaaat	aagcagtgga	gtgggggata	gtgagagtga	aggcctgaac	960
agccctgtca	aagttgctcg	aaagcgggaag	agaatggtga	ctggaaatgg	ctctcttaaa	1020
aggaaaagct	ctaggaagga	aacgccctca	gccaccaaac	aagcaactag	catttcatca	1080
gaaaccaaga	atactttgag	agctttctct	gcccctcaaa	attctgaatc	ccaagcccac	1140
gttagtggag	gtggtgatga	cagtagtcgc	cctactgttt	ggtatcatga	aactttagaa	1200
tggcttaagg	agggaaaagag	aagagatgag	cacaggagga	ggcctgatca	ccccgatttt	1260
gatgcattca	cactctatgt	gcctgaggat	ttcctcaatt	cttgactacc	tgggatgagg	1320
aagtgggtggc	agattaagtc	tcagaacttt	gatcttgtca	tctgttataa	ggtggggaaa	1380
ttttatgagc	tgtaccacat	ggatgctctt	attggagtca	gtgaactggg	gctgggtattc	1440
atgaaaggca	actgggcccc	ttctggcttt	cctgaaattg	catttggccg	ttattcagat	1500
tccctgggtgc	agaagggcta	taaagtagca	cgagtggaac	agactgagac	tccagaaatg	1560
atggaggcac	gatgtagaaa	gatggcacat	atatccaagt	atgatagagt	ggtgaggagg	1620
gagatctgta	ggatcattac	caagggtaca	cagacttaca	gtgtgctgga	aggtgatccc	1680
tctgagaact	acagtaagta	tcttcttagc	ctcaaagaaa	aagaggaaga	ttcttctggc	1740
catactcgtg	catatggtgt	gtgctttgtt	gatacttcac	tgggaaagtt	tttcataggt	1800
cagttttcag	atgatcgcca	ttgttcgaga	tttaggactc	tagtggcaca	ctatccccc	1860
gtacaagttt	tatttgaaaa	aggaaatctc	tcaaaggaaa	ctaaaacaat	tctaaagagt	1920

tcattgtcct	gttctcttca	ggaaggtctg	ataccggct	cccagttttg	ggatgcatcc	1980
aaaactttga	gaactctcct	tgaggaagaa	tattttaggg	aaaagctaag	tgatggcatt	2040
ggggtgatgt	taccccaggt	gcttaaagg	atgacttcag	agtctgattc	cattggggtg	2100
acaccaggag	agaaaagtga	attggccctc	tctgctctag	gtggttgtgt	cttctacctc	2160
aaaaaatgcc	ttattgatca	ggagctttta	tcaatggcta	atattgaaga	atatattccc	2220
ttggattctg	acacagtcag	cactacaaga	tctgggtgcta	tcttcaccaa	agcctatcaa	2280
cgaatgggtgc	tagatgcagt	gacattaaac	aacttggaga	tttttctgaa	tggaacaaat	2340
ggttctactg	aaggaaccct	actagagagg	gttgatactt	gccatactcc	ttttggtaag	2400
cggctcctaa	agcaatggct	ttgtgcccc	ctctgtaacc	attatgctat	taatgatcgt	2460
ctagatgcc	tagaagacct	catggttgtg	cctgacaaaa	tctccgaagt	tgtagagctt	2520
ctaaagaagc	ttccagatct	tgagaggcta	ctcagtaaaa	ttcataatgt	tgggtctccc	2580
ctgaagagtc	agaaccaccc	agacagcagg	gctataatgt	atgaagaaac	tacatacagc	2640
aagaagaaga	ttattgattt	tctttctgct	ctggaaggat	tcaaagtaat	gtgtaaaatt	2700
atagggatca	tggaagaagt	tgctgatgg	tttaagtcta	aaatccttaa	gcaggtcatt	2760
tctctgcaga	caaaaaatcc	tgaaggtcgt	tttcctgatt	tgactgtaga	attgaaccga	2820
tgggatacag	cctttgacca	tgaaaaggct	cgaaagactg	gacttattac	tcccaaagca	2880
ggctttgact	ctgattatga	ccaagctctt	gctgacataa	gagaaaatga	acagagcctc	2940
ctggaatacc	tagagaaaca	gcgcaacaga	attggctgta	ggaccatagt	ctattggggg	3000
attggtagga	accggtacca	gctggaaatt	cctgagaatt	tcaccactcg	caatttgcca	3060
gaagaatacg	agttgaaatc	taccaagaag	ggctgtaaac	gatactggac	caaaactatt	3120
gaaaagaagt	tggctaattct	cataaatgct	gaagaacgga	gggatgtatc	attgaaggac	3180
tgcatgcggc	gactgttcta	taactttgat	aaaaattaca	aggactggca	gtctgctgta	3240
gagtgtatcg	cagtgttgg	tgttttactg	tgcttggtg	actatagtcg	agggggtgat	3300
ggctctatgt	gtcggccagt	aattctgttg	ccggaagata	ccccccctt	cttagagctt	3360
aaaggatcac	gccatccttg	cattacgaag	actttttttg	gagatgattt	tattcctaatt	3420
gacattctaa	taggctgtga	ggaagaggag	caggaaaatg	gcaaagccta	ttgtgtgctt	3480
gttactggac	caaatatggg	gggcaagtct	acgcttatga	gacaggctgg	cttattagct	3540
gtaatggccc	agatgggttg	ttacgtccct	gctgaagtgt	gcaggctcac	accaattgat	3600
agagtgttta	ctagacttgg	tgcttcagac	agaataatgt	cagggtgaaag	tacatttttt	3660
gttgaattaa	gtgaaactgc	cagcatactc	atgcatgcaa	cagcacattc	tctgggtgctt	3720
gtggatgaat	taggaagagg	tactgcaaca	tttgatggga	cggcaatagc	aaatgcagtt	3780
gttaaagaac	ttgctgagac	tataaaatgt	cgtacattat	tttcaactca	ctaccattca	3840
ttagtagaag	attattctca	aaatgttgct	gtgcgcctag	gacatatggc	atgcatggta	3900
gaaaatgaat	gtgaagaccc	cagccaggag	actattacgt	tcctctataa	attcattaag	3960

ggagcttgtc ctaaaagcta tggctttaat gcagcaaggc ttgctaattct cccagaggaa 4020  
gttattcaaa agggacatag aaaagcaaga gaatttgaga agatgaatca gtcactacga 4080  
ttatttcggg aagtttgctt ggctagtgaagggtcaactg tagatgctga agctgtccat 4140  
aaattgctga ctttgattaa ggaattatag actgactaca ttggaagctt tgagttgact 4200  
tctgaccaa ggtggtaaat tcagacaaca ttatgatcta ataaacttta ttttttaaaa 4260  
atga 4264